PATENT COOPERATION TREATY

PCT

NOTIFICATION OF ELECTION

(PCT Rule 61.2)

From the INTERNATIONAL BUREAU

To:

United States Patent and Trademark Office
(Box PCT)
Crystal Plaza 2
Washington, DC 20231
ETATS-UNIS D'AMERIQUE

in its capacity as elected Office

Date of mailing (day/month/year)
31 January 1997 (31.01.97)

International application No.	Applicant's or agent's file reference
PCT/GB96/01409	GWS/18446

International filing date (day/month/year)	Priority date (day/month/year)
12 June 1996 (12.06.96)	12 June 1995 (12.06.95)

Applicant

ELMORE, Michael, James et al

1. The designated Office is hereby notified of its election made:

   [X] in the demand filed with the International Preliminary Examining Authority on:

   13 January 1997 (13.01.97)

   [ ] in a notice effecting later election filed with the International Bureau on:

2. The election [X] was

   [ ] was not

   made before the expiration of 19 months from the priority date or, where Rule 32 applies, within the time limit under Rule 32.2(b).

The International Bureau of WIPO 34, chemin des Colombettes 1211 Geneva 20, Switzerland	Authorized officer F. Gateau
Facsimile No.: (41-22) 740.14.35	Telephone No.: (41-22) 730.91.11

Form PCT/IB/331 (July 1992)   1381127

PATENT COOPERATION TREATY

PCT

NOTIFICATION OF THE RECORDING OF A CHANGE

(PCT Rule 92bis.1 and
Administrative Instructions, Section 422)

From the INTERNATIONAL BUREAU

To:

SCHLICH, George, William
Mathys & Squire
100 Gray's Inn Road
London WC1X 8AL
ROYAUME-UNI

Date of mailing (day/month/year) 27 March 1997 (27.03.97)	
Applicant's or agent's file reference GWS/18446	**IMPORTANT NOTIFICATION**
International application No. PCT/GB96/01409	International filing date (day/month/year) 12 June 1996 (12.06.96)

---

1. The following indications appeared on record concerning:

   [X] the applicant   [X] the inventor   [ ] the agent   [ ] the common representative

Name and Address	State of Nationality	State of Residence
	Telephone No.	
	Facsimile No.	
	Teleprinter No.	

---

2. The International Bureau hereby notifies the applicant that the following change has been recorded concerning:

   [ ] the person   [ ] the name   [ ] the address   [ ] the nationality   [ ] the residence

Name and Address	State of Nationality GB	State of Residence GB
TITBALL, Richard, William Chemical & Biological Sector Porton Down Salisbury, Wiltshire SP4 0JQ United Kingdom	Telephone No.	
	Facsimile No.	
	Teleprinter No.	

---

3. Further observations, if necessary:
   Additional applicant/inventor for the purposes of US only.

---

4. A copy of this notification has been sent to:

   [X] the receiving Office   [ ] the designated Offices concerned

   [ ] the International Searching Authority   [X] the elected Offices concerned

   [X] the International Preliminary Examining Authority   [ ] other:

The International Bureau of WIPO 34, chemin des Colombettes 1211 Geneva 20, Switzerland	Authorized officer Marie-José Devillard
Facsimile No.: (41-22) 740.14.35	Telephone No.: (41-22) 730.91.11

Form PCT/IB/306 (March 1994)   001457953

5

PATENT COOPERATION TREATY

PCT

INTERNATIONAL PRELIMINARY EXAMINATION REPORT

(PCT Article 36 and Rule 70)

Applicant's or agent's file reference GWS/18446	FOR FURTHER ACTION	See Notification of Transmittal of International Preliminary Examination Report (Form PCT/IPEA/416)
International application No. PCT/GB 96/ 01409	International filing date *(day/month/year)* 12/06/1996	Priority date *(day/month/year)* 12/06/1995

International Patent Classification (IPC) or national classification and IPC
C12N15/31

Applicant
MICROBIOLOGICAL RESEARCH AUTHORITY et al.

1. This international preliminary examination report has been prepared by this International Preliminary Examining Authority and is transmitted to the applicant according to Article 36.

2. This **REPORT** consists of a total of ___7___ sheets, including this cover sheet.

   ☒ This report is also accompanied by ANNEXES, i.e., sheets of the description, claims and/or drawings which have been amended and are the basis for this report and/or sheets containing rectifications made before this Authority (see Rule 70.16 and Section 607 of the Administrative Instructions under the PCT).

   These annexes consists of a total of ___4___ sheets.

3. This report contains indications and corresponding pages relating to the following items:

   I ☒ Basis of the report

   II ☐ Priority

   III ☐ Non-establishment of opinion with regard to novelty, inventive step and industrial applicability

   IV ☐ Lack of unity of invention

   V ☒ Reasoned statement under Article 35(2) with regard to novelty, inventive step or industrial applicability; citations and explanations supporting such statement

   VI ☐ Certain documents cited

   VII ☒ Certain defects in the international application

   VIII ☒ Certain observations on the international application

Date of submission of the demand 13/01/1997	Date of completion of this report **2 6. 09. 97**
Name and mailing address of the IPEA/ European Patent Office D-80298 Munich Tel. (+49-89) 2399-0, Tx: 523656 epmu d Fax: (+49-89) 2399-4465	Authorized officer **V. Kaas** Telephone No.

I. Basis of the report

1. This report has been drawn up on the basis of (Replacement sheets which have been furnished to the receiving Office in response to an invitation under Article 14 are referred to in this report as "originally filed" and are not annexed to the report since they do not contain amendments.):

[ ] the international application as originally filed.

[x] the description, pages 1-14_____, as originally filed,

  pages _____, filed with the demand,

  pages _____, filed with the letter of _____,

  pages _____, filed with the letter of _____,

[x] the claims, Nos. _____, as originally filed,

  Nos. _____, as amended under Article 19,

  Nos. _____, filed with the demand,

  Nos. 1-23_____, filed with the letter of 15/08/97,

  Nos. _____, filed with the letter of _____,

[x] the drawings, sheets/fig 1/4-4/4_____, as originally filed,

  sheets/fig _____, filed with the demand,

  sheets/fig _____, filed with the letter of _____,

  sheets/fig _____, filed with the letter of _____.

2. The amendments have resulted in the cancellation of:

  [ ] the description, pages _____.

  [ ] the claims,   Nos. _____.

  [ ] the drawings,   sheets/fig _____.

3. [ ] This report has been established as if (some of) the amendments had not been made, since they have been considered to go beyond the disclosure as filed (Rule 70.2(c)):

4. Additional observations, if necessary:

Form PCT/IPEA/409 (sheet 1) (January 1994)

V. Reasoned statement under Article 35(2) with regard to novelty, inventive step and industrial applicability; citations and explanations supporting such statement

1. STATEMENT

Novelty (N)	Claims 1-23	YES
	Claims	NO
Inventive Step (IS)	Claims 1-23	YES
	Claims	NO
Industrial Applicability (IA)	Claims 1-21	YES
	Claims 22-23 (no assessment)	NO

2. CITATIONS AND EXPLANATIONS

1) Reference is made to the following documents:

D1 : Database Medline, File Server STN Karlsruhe, Abstract 77064466

D2 : Abstracts of the 95th General Meeting of the American Society for Microbiology, 21-25 May 1995, Washington D.C USA, page 289, Abstract E-49.

D3 : Biochemistry, vol. 33, 1994, pages 7014-7020

D4 : WO-A-94/03615

2) The present application appears to satisfy the criterions set forth in Articles 33(2) and 33(3) PCT in respect of the prior art as defined in the regulations (Rule 64(1)-(3) PCT).

D1 discloses the purification of the toxin from Clostridium botulinum type F. However, D1 teaches immunization of guinea pigs with the toxoid obtained from said toxin. The vaccinated guinea pigs are resistent to

challenge by type F botulinal toxin. There is neither a disclosure nor suggestion of an immunization with a polypeptide free of toxoid.

D2 discloses that the vaccination of humans with pure type F toxoid lead to the production of a protective serum antibody response. Here again, D2 is silent as to an immunization with a polypeptide which is free of toxoid.

D3 discloses the preparation by recombinant techniques of a maltose binding protein (MBP)- tetanus toxin light chain (LC) fusion product, wherein residue Glu234 of the light chain is replaced by Ala, MBP being used as an affinity tag to facilitate purification and subsequent isolation of free LC. It is shown that the Ala234-LC mutant lacks neurotoxicity in mice. However, it is clear that the application of the specific amino acid replacement taught by D3 is restricted to the tetanus toxin.

D4 discloses the preparation of a fusion protein consisting of a polypeptide immunogen linked by a hinge fragment to an antigenic determinant of a pathogen organism. As a preferred embodiment, D4 discloses as polypeptide immunogen the Tetanus toxin C fragment and as antigenic sequence, the Schistosoma mansoni P28 protein, the latter being capable of being affinity purified on a glutathione agarose matrix (see page 10, line 6--page 11, paragraph 2). The use of the fusion protein as a vaccine is also disclosed (see page 13, third paragraph- page 15, first paragraph; Example 5). It is not derivable from D4 whether the Tetanus toxin C fragment used in D4 is free of toxin activity, which is one of the requirements recited in the present claims. Moreover, this requirement is not mentioned in D4 as regards the alternatives listed therein on page 5, lines 10-16. It therefore appears that even if the skilled man would have been tempted to apply the teaching of D4 to the

preparation of a fusion protein capable of protecting against a type F botulinum toxin, he would not have arrived at the claimed subject-matter, the above requirement being not obviously derivable from D4.

Claims 1-23 therefore appear to be novel and to imply an inventive step according to Articles 33(2) and 33(3) PCT.

4)   The priority documents pertaining to the present application were not available at the time of establishing this preliminary examination report. Hence, it is based on the assumption that all claims enjoy priority rights from the filing date of the priority document. The document "Biochemistry, vol. 34, Nov. 1995, pages 15175-15181" cited in the international search report has therefore not been considered to be part of the prior art as defined in the regulations (Rule 64(1)-(3) PCT).

5)   Claims 1-21 are susceptible of industrial applicability as defined in Article 33(4) PCT.

6)   For the assessment of present claims 22 and 23 on the question as to whether they are industrially applicable, no unified criteria exist in the PCT. The patentability can also be dependent upon the formulation of the claims. The EPO, for example, does not recognize as industrially applicable the subject-matter of claims to the use of a compound in medical treatment, but will allow, however, claims to a known compound for first use in medical treatment and the use of such a compound for the manufacture of a medicament for a new medical treatment.

VII. Certain defects in the international application

The following defects in the form or contents of the international application have been noted:

1) Contrary to the requirements of Rule 5.1(a)(ii) PCT, the relevant background art disclosed in document D4 is not mentioned in the description, nor is this document identified therein.

2) The description is not in conformity with the claims as required by Rule 5.1(a)(iii) PCT.

VIII. Certain observations on the international application

The following observations on the clarity of the claims, description, and drawings or on the question whether the claims are fully supported by the description, are made:

1) Claims 1, 2, 7, 17 and 19 do not meet the requirements of Article 6 PCT in that the matter for which protection is sought is defined in terms of the result to be achieved. In this instance, however, such a formulation is not allowable because it seems possible to define the subject-matter in more concrete terms, viz. in terms of how the result is to be achieved. In this respect, the only substantial support which can be found in the description for these claims is represented by the amino sequences SEQ ID NO:1 to SEQ ID NO:4.

2) Claims 7, 17 and 19 also lack clarity (Article 6 PCT) in that the essential technical feature that the polypeptide is free of toxoid (see page 3, line 19; page 4, last line) is not recited in the claims.

3) Claim 14 lacks clarity (Article 6 PCT) in that it does not recite the essential technical feature that the chromatography column is an affinity chromatography column (see page 8, line 15).

INTERNATIONAL SEARCH REPORT

A. CLASSIFICATION OF SUBJECT MATTER
IPC 6   C12N15/31   C12N15/62   C07K14/33   A61K39/08

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

IPC 6   C07K

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practical, search terms used)

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	DATABASE MEDLINE FILE SERVER STN KARLSRUHE ABSTRACT 77064466, HATHEWAY: "TOXOID OF CLOSTRIDIUM BOTULINUM TYPE F: PURIFICATION AND IMMUNOGENICITY STUDIES" XP002015940 & APPLIED AND ENVIROMENTAL MICROBIOLOGY, (1976 FEB) 31 (2) 234-42. see abstract ---	1,2,12, 22
X	ABSTRACTS OF THE 95TH GENERAL MEETING OF THE AMERICAN SOCIETY FOR MICROBIOLOGY, 21 - 25 May 1995, WASHINGTON D.C,USA, page 289 XP002015937 MONTGOMERY ET AL: "EVALUATION OF BOTULINUM TYPE F VACCINE BY ELISA" see abstract E-49 ---	1,2,12, 22

-/--

[X] Further documents are listed in the continuation of box C.   [X] Patent family members are listed in annex.

* Special categories of cited documents :

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier document but published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.

"&" document member of the same patent family

Date of the actual completion of the international search	Date of mailing of the international search report
16 October 1996	0 4. 11. 96
Name and mailing address of the ISA European Patent Office, P.B. 5818 Patentlaan 2 NL - 2280 HV Rijswijk Tel. (+31-70) 340-2040, Tx. 31 651 epo nl, Fax (+31-70) 340-3016	Authorized officer Sitch, W

page 1 of 3

C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT

Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	BIOCHEMISTRY, vol. 33, 1994, pages 7014-7020, XP002015938 LI ET AL: "A SINGLE MUTATION IN THE RECOMBINANT LIGHT CHAIN OF TETANUS TOXIN ABOLISHES ITS PROTEOLYTIC ACTIVITY AND REMOVES THE TOXICITY SEEN AFTER RECONSTITUTION WITH NATIVE HEAVY CHAIN" see the whole document ---	7,8,10, 12,13, 17-19,21
X	WO,A,94 03615 (MEDEVA HOLDINGS B V ;KHAN MOHAMMED ANJAM (GB); HORMAECHE CARLOS ES) 17 February 1994 see page 10, paragraph 2 - page 11, paragraph 2 ---	7,8,10, 12,13, 17-19,21
X	DATABASE CHEMICAL ABSTRACTS FILE SERVER STN KARLSRUHE ABSTRACT 123:219673, MINTON: "PHYSICAL CHARACTERIZATION OF CLOSTRIDIUM BOTULINUM NEUROTOXIN GENES" XP002015941 & REPORT (1993),ORDER NO.AD-A272939 ... see abstract ---	7,8,10, 12,13, 17-19, 21-23
A	DATABASE CHEMICAL ABSTRACTS FILE SERVER STN KARLSRUHE ABSTRACT 119:21962, MINTON: "PHYSICAL CHARACTERIZATION OF CLOSTRIDIUM BOTULINUM NEUROTOXIN GENES" XP002015942 & REPORT (1992),ORDER NO.AD-A248904... see abstract ---	1-23
A	SYSTEM.APPL.MICROBIOL., vol. 18, no. 1, May 1995, pages 23-31, XP000605363 ELMORE ET AL: "NUCLEOTIDE SEQUENCE OF THE GENE CODING FOR PROTEOLYTIC (GROUP I) CLOSTRIDIUM BOTULINUM TYPE F NEUROTOXIN: GENEALOGICAL COMPARISON WITH OTHER CLOSTRIDIAL NEUROTOXINS" see the whole document ---	1-23
A	WO,A,94 21684 (PHLSB ;NIBSC (GB); SESARDIC DOROTHEA (GB); CHAN WOON LING (GB); SH) 29 September 1994 see page 12; table 1 see page 17,last paragraph-page 18,last paragraph see page 20; table 6 ---	1-23

-/--

C.(Continuation) D CUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	TRENDS IN BIOTECHNOLOGY, vol. 8, no. 4, April 1990, pages 88-93, XP000103110 SASSENFELD: "ENGINEERING PROTEINS FOR PURIFICATION" see the whole document ---	7,8,10, 12,13, 17-19, 21-23
P,X	BIOCHEMISTRY, vol. 34, November 1995, pages 15175-15181, XP002015939 ZHOU ET AL: "EXPRESSION AND PURIFICATION OF THE LIGHT CHAIN OF BOTULINUM NEUROTOXIN A: A SINGLE MUTATION ABOLISHES ITS CLEAVAGE OF SNAP-25 AND NEUROTOXICITY AFTER RECONSTITUTION WITH THE HEAVY CHAIN" see the whole document -----	7,8,10, 12,13, 17-19, 21-23

2

INTERNATIONAL SEARCH REPORT

(PCT Article 18 and Rules 43 and 44)

Applicant's or agent's file reference GWS/18446	**FOR FURTHER** **ACTION**	see Notification of Transmittal of International Search Report (Form PCT/ISA/220) as well as, where applicable, item 5 below.	
International application No. PCT/GB 96/01409	International filing date*(day/month/year)* 12/06/1996	(Earliest) Priority Date *(day/month/year)* 12/06/1995	

Applicant MICROBIOLOGICAL RESEARCH AUTHORITY et al.

This International Search Report has been prepared by this International Searching Authority and is transmitted to the applicant according to Article 18. A copy is being transmitted to the International Bureau.

This International Search Report consists of a total of ___5___ sheets.

[X] It is also accompanied by a copy of each prior art document cited in this report.

1. [X] Certain claims were found unsearchable (see Box I).

2. [ ] Unity of invention is lacking (see Box II).

3. [X] The international application contains disclosure of a **nucleotide and/or amino acid sequence listing** and the international search was carried out on the basis of the sequence listing

   [X] filed with the international application.

   [ ] furnished by the applicant separately from the international application,

     [ ] but not accompanied by a statement to the effect that it did not include matter going beyond the disclosure in the international application as filed.

   [ ] Transcribed by this Authority

4. With regard to the **title**,   [ ] the text is approved as submitted by the applicant.

     [X] the text has been established by this Authority to read as follows:

   Type F Botulinum toxin and use thereof

5. With regard to the **abstract**,

   [X] the text is approved as submitted by the applicant.

   [ ] the text has been established, according to Rule 38.2(b), by this Authority as it appears in Box III. The applicant may, within one month from the date of mailing of this International Search Report, submit comments to this Authority.

6. The figure of the **drawings** to be published with the abstract is:

   Figure No. ___4___   [X] as suggested by the applicant.   [ ] None of the figures.

             [ ] because the applicant failed to suggest a figure.

             [ ] because this figure better characterizes the invention.

INTERNATIONAL SEARCH REPORT

Box I Observations where certain claims were found unsearchable (Continuation of item 1 of first sheet)

This International Search Report has not been established in respect of certain claims under Article 17(2)(a) for the following reasons:

1. [X] Claims Nos.:   22,23
because they relate to subject matter not required to be searched by this Authority, namely:

> Remark: Although claims 22 and 23 are directed to a method of treatment of the human/animal body, the search has been carried out and based on the alleged effects of the compound/composition.

2. [ ] Claims Nos.:
because they relate to parts of the International Application that do not comply with the prescribed requirements to such an extent that no meaningful International Search can be carried out, specifically:

3. [ ] Claims Nos.:   .
because they are dependent claims and are not drafted in accordance with the second and third sentences of Rule 6.4(a).

Box II Observations where unity of invention is lacking (Continuation of item 2 of first sheet)

This International Searching Authority found multiple inventions in this international application, as follows:

1. [ ] As all required additional search fees were timely paid by the applicant, this International Search Report covers all searchable claims.

2. [ ] As all searchable claims could be searches without effort justifying an additional fee, this Authority did not invite payment of any additional fee.

3. [ ] As only some of the required additional search fees were timely paid by the applicant, this International Search Report covers only those claims for which fees were paid, specifically claims Nos.:

4. [ ] No required additional search fees were timely paid by the applicant. Consequently, this International Search Report is restricted to the invention first mentioned in the claims; it is covered by claims Nos.:

Remark on Protest
     [ ] The additional search fees were accompanied by the applicant's protest.

     [ ] No protest accompanied the payment of additional search fees.

Form PCT/ISA/210 (continuation of first sheet (1)) (July 1992)

A. CLASSIFICATION OF SUBJECT MATTER
IPC 6   C12N15/31   C12N15/62   C07K14/33   A61K39/08

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)
IPC 6   C07K

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practical, search terms used)

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	DATABASE MEDLINE FILE SERVER STN KARLSRUHE ABSTRACT 77064466, HATHEWAY: "TOXOID OF CLOSTRIDIUM BOTULINUM TYPE F: PURIFICATION AND IMMUNOGENICITY STUDIES" XP002015940 & APPLIED AND ENVIROMENTAL MICROBIOLOGY, (1976 FEB) 31 (2) 234-42. see abstract ---	1,2,12, 22
X	ABSTRACTS OF THE 95TH GENERAL MEETING OF THE AMERICAN SOCIETY FOR MICROBIOLOGY, 21 - 25 May 1995, WASHINGTON D.C,USA, page 289 XP002015937 MONTGOMERY ET AL: "EVALUATION OF BOTULINUM TYPE F VACCINE BY ELISA" see abstract E-49 ---	1,2,12, 22

-/--

| X | Further documents are listed in the continuation of box C. | | X | Patent family members are listed in annex. |

* Special categories of cited documents :

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier document but published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art

"&" document member of the same patent family

Date of the actual completion of the international search	Date of mailing of the international search report
16 October 1996	0 4. 11. 96

| Name and mailing address of the ISA
European Patent Office, P.B. 5818 Patentlaan 2
NL - 2280 HV Rijswijk
Tel. (+31-70) 340-2040, Tx. 31 651 epo nl,
Fax: (+31-70) 340-3016 | Authorized officer

Sitch, W |

Form PCT/ISA/210 (second sheet) (July 1992)

International Application No
PCT/GB 96/01409

C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	BIOCHEMISTRY, vol. 33, 1994, pages 7014-7020, XP002015938 LI ET AL: "A SINGLE MUTATION IN THE RECOMBINANT LIGHT CHAIN OF TETANUS TOXIN ABOLISHES ITS PROTEOLYTIC ACTIVITY AND REMOVES THE TOXICITY SEEN AFTER RECONSTITUTION WITH NATIVE HEAVY CHAIN" see the whole document ---	7,8,10, 12,13, 17-19,21
X	WO,A,94 03615 (MEDEVA HOLDINGS B V ;KHAN MOHAMMED ANJAM (GB); HORMAECHE CARLOS ES) 17 February 1994 see page 10, paragraph 2 - page 11, paragraph 2 ---	7,8,10, 12,13, 17-19,21
X	DATABASE CHEMICAL ABSTRACTS FILE SERVER STN KARLSRUHE ABSTRACT 123:219673, MINTON: "PHYSICAL CHARACTERIZATION OF CLOSTRIDIUM BOTULINUM NEUROTOXIN GENES" XP002015941 & REPORT (1993),ORDER NO.AD-A272939 ... see abstract ---	7,8,10, 12,13, 17-19, 21-23
A	DATABASE CHEMICAL ABSTRACTS FILE SERVER STN KARLSRUHE ABSTRACT 119:21962, MINTON: "PHYSICAL CHARACTERIZATION OF CLOSTRIDIUM BOTULINUM NEUROTOXIN GENES" XP002015942 & REPORT (1992),ORDER NO.AD-A248904... see abstract ---	1-23
A	SYSTEM.APPL.MICROBIOL., vol. 18, no. 1, May 1995, pages 23-31, XP000605363 ELMORE ET AL: "NUCLEOTIDE SEQUENCE OF THE GENE CODING FOR PROTEOLYTIC (GROUP I) CLOSTRIDIUM BOTULINUM TYPE F NEUROTOXIN: GENEALOGICAL COMPARISON WITH OTHER CLOSTRIDIAL NEUROTOXINS" see the whole document ---	1-23
A	WO,A,94 21684 (PHLSB ;NIBSC (GB); SESARDIC DOROTHEA (GB); CHAN WOON LING (GB); SH) 29 September 1994 see page 12; table 1 see page 17,last paragraph-page 18,last paragraph see page 20; table 6 ---	1-23

-/--

Form PCT/ISA/210 (continuation of second sheet) (July 1992)

C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	TRENDS IN BIOTECHNOLOGY, vol. 8, no. 4, April 1990, pages 88-93, XP000103110 SASSENFELD: "ENGINEERING PROTEINS FOR PURIFICATION" see the whole document ---	7,8,10, 12,13, 17-19, 21-23
P,X	BIOCHEMISTRY, vol. 34, November 1995, pages 15175-15181, XP002015939 ZHOU ET AL: "EXPRESSION AND PURIFICATION OF THE LIGHT CHAIN OF BOTULINUM NEUROTOXIN A: A SINGLE MUTATION ABOLISHES ITS CLEAVAGE OF SNAP-25 AND NEUROTOXICITY AFTER RECONSTITUTION WITH THE HEAVY CHAIN" see the whole document -----	7,8,10, 12,13, 17-19, 21-23

2

Form PCT/ISA/210 (continuation of second sheet) (July 1992)

Patent document cited in search report	Publication date	Patent family member(s)		Publication date
WO-A-9403615	17-02-94	AU-A-	4719393	03-03-94
		CA-A-	2141427	17-02-94
		EP-A-	0652962	17-05-95
		FI-A-	950396	30-01-95
		JP-T-	8503602	23-04-96
		NO-A-	950348	28-03-95
WO-A-9421684	29-09-94	AU-A-	6432594	11-10-94
		CA-A-	2158748	29-09-94
		EP-A-	0690875	10-01-96

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 6 :		(11) International Publicati n Number:	WO 96/41881
C12N 15/31, 15/62, C07K 14/33, A61K 39/08	A1	(43) International Publication Date:	27 December 1996 (27.12.96)

(21) International Application Number: PCT/GB96/01409

(22) International Filing Date: 12 June 1996 (12.06.96)

(30) Priority Data:
9511909.5   12 June 1995 (12.06.95)   GB

(71) Applicant (for all designated States except US): MICROBIO-LOGICAL RESEARCH AUTHORITY [GB/GB]; CAMR (Centre For Applied Microbiology & Research), Porton Down, Salisbury, Wiltshire SP4 0JG (GB).

(72) Inventors; and
(75) Inventors/Applicants (for US only): ELMORE, Michael, James [GB/GB]; 8 St Mary's Court, Eastrop Lane, Basingstoke, Hants RG21 4AT (GB). MAUCHLINE, Margaret, Lamble [GB/GB]; Three Dormers, Bakers' Hill, Semley, Shaftesbury, Dorset SP7 9BQ (GB). MINTON, Nigel, Peter [GB/GB]; 27 Moberly Road, Salisbury, Wiltshire SP1 3BZ (GB). PASECHNIK, Vladimir, Artymovich [GB/GB]; 1 Copper Beech Close, Shrewton, Wiltshire SP4 4HU (GB).

(74) Agent: SCHLICH, George, William; Mathys & Squire, 100 Gray's Inn Road, London WC1X 8AL (GB).

(81) Designated States: AU, CA, JP, US, European patent (AT, BE, CH, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE).

Published
    With international search report.
    Before the expiration of the time limit for amending the claims and to be republished in the event of the receipt of amendments.

(54) Title: TYPE F BOTULINUM TOXIN AND USE THEREOF

ANTIGENICITY OF SERUM AFTER IMMUNISATION OF MICE WITH
MBP-BoNT/F (868-1278) RECOMBINANT PROTEIN
ANTIGEN: BoNT/F, 13ng/WELL
SERA S-2 (1) AND S-3 (2) WERE AFTER SECOND AND THIRD BOOSTS

SERUM S-2, S-3 AS WELL AS NON-IMMUNE SERUM WERE FIRST DILUTED 1:50 AND 1:3
AT EACH NEXT STEP

(3): NON-IMMUNE SERA

(57) Abstract

    A polypeptide free of toxin activity gives protection against botulinum type F toxin. A fusion protein of a fragment of a toxin molecule and a purification moeity enables purification of the fragment from solution. Pharmaceutical compositions containing the polypeptide and the fusion protein are described.

Type F Botulinum toxin and use thereof

The present invention relates to type F botulinum toxin, to a fragment of type F botulinum neurotoxin, to production of the fragment by recombinant means and to a synthetic gene encoding the fragment. In particular, the invention relates to a novel polypeptide fragment capable of eliciting an immunological response that is protective against type F botulinum neurotoxin (BoNT/F) in man or animals and to a vaccine containing the fragment.

Botulinum neurotoxins (BoNTs) are high molecular weight proteins (approx. 150,000 Da) which exert potent neuroparalytic effects on vertebrates. They are elaborated by anaerobic Gram-positive bacteria belonging to the genus *Clostridium*. The majority of clostridia which produce BoNT are classified as *Clostridium botulinum*. In recent years, however, isolates which resemble *Clostridium barati* and *Clostridium butyricum* have been shown to produce BoNT. On the basis of antigenicity, BoNT has been subdivided into seven distinct types, designated A to G. All seven neurotoxins (BoNT/A to BoNT/G) are synthesised as a single-chain 150,000 Da molecule which subsequently become nicked to the more potent di-chain form, composed of a heavy (H) chain (approx. 100,000 Da) and a light (L) chain (approx. 50,000 Da) linked by at least one disulphide bridge.

The action of BoNT involves three distinct phases. In the first phase the toxins become bound to acceptors on the external surface of the targeted neural cells. This is followed by an energy dependent internalisation step in which the toxin, or part of it, enters the cell. Thereafter, the active moiety of the toxin causes nerve cell dysfunction by blocking the intracellular release of the neurotransmitter, acetylcholine, at the nerve periphery, causing flaccid paralysis. The L chain possesses the catalytic activity responsible for cell poisoning and the H chain delivers this moiety to the cell cytoplasm by mediating binding of the toxin to the cell and subsequent internalisation.

The entire amino acid sequences of all 7 BoNTs are now known (Minton, N.P. (1995). Current Topics in Microbiology and Immunology 195: 161–187), revealing them to be surprisingly divergent in their primary amino acid sequences. Thus, sequence identity amongst the different serotypes generally does not exceed 40%, with those areas of homology localised to discrete domains which are interspersed with amino acid tracts exhibiting little overall similarity. Between the different L chains (average size 439), 63 amino acids are absolutely conserved. Throughout the H chains (average size 843) 97 amino acids are identical. The most notable areas of conservation include:– the two cysteine residues involved in the disulphide bond formation between the L and H chain; the histidine rich motif within the L chain associated with metalloprotease activity; and a highly conserved PYI/VXALN–motif found adjacent to regions identified as possessing membrane spanning potential. The most notable tract of sequence divergence amongst toxins is localised to the COOH–terminus of their respective H chains (amino acid 1124 onwards of BoNT/A). This would appear to be consistent with the notion that this domain is involved in neurotoxin binding and that different toxins target different acceptors on neural cell surfaces.

The effectiveness of modern food–preserving processes in Western countries has made outbreaks of botulism extremely rare. The frequent use of C.botulinum as a test organism in the food industry, and the growing use of the toxin by neurobiochemists, has, however, increased the need for human vaccines. The formulation of these vaccines has changed little since the early 1950s: partially purified preparations of the neurotoxins are toxoided by formaldehyde treatment and absorbed onto precipitated aluminium salts. Using such methodology, polyvalent vaccines (against ABCDE or ABEF) for human immunisation are currently available. Such vaccines suffer from the drawback of low immune response and considerable batch to batch variation due to the high proportion (60–90%) of contaminating proteins in toxoid preparations. Recent work has therefore concentrated on the development of procedures for the purification of toxins to near–homogeneity. The use of purified toxins in the production of vaccines, however, suffers from the drawbacks, first, of having to produce them under high containment and, secondly, of requiring the presence of low levels of formaldehyde to prevent possible reversion of the toxoid to the active state.

Production of subunit vaccines against other organisms and toxins has been investigated by a number of laboratories. This work has focused on the best known toxin subtypes, namely A and B, leading to new vaccines giving specific immunity against toxins of type A or B. Each new vaccine, however, may not give protection against other toxin subtypes.

Recombinant production of vaccine components has brought great advances in vaccine purity and volume of production. A.J. Makoff et al, in Bio/Technology, volume 7, October 1989, pages 1043-1046, describe the expression of a tetanus toxin fragment in *E.coli,* and its purification and potential use as a vaccine. The technique described nevertheless requires a large number of steps to recover purified vaccine components from the host cells.

It is an object of this invention to produce a vaccine against a type F botulinum toxin. It is another object to simplify vaccine manufacture. A further object is to improve production of botulinum toxin vaccines. A still further object of the invention is to overcome or at least mitigate problems and/or limitations in existing vaccines and methods of production.

According to a first aspect of the invention there is provided a polypeptide free of botulinum toxin activity which induces protective immunity to a type F botulinum toxin. The polypeptide is useful in manufacture of a vaccine against type F toxin, and in contrast to prior art compositions such as polyvalent vaccines is not a toxoid and does not need pretreatment with formaldehyde. Also in contrast to prior art compositions the polypeptide is generally of smaller size than the toxin itself.

An embodiment of the first aspect of the invention provides a polypeptide characterized in that it:-

    (a)   is free of botulinum toxin activity, and

    (b)   is capable of eliciting, in a mammal, an immunological response that is protective against type F botulinum toxin.

– 4 –

The term "protective" used in conjunction with "immunity" and "immunological response" is used to indicate an increased ability to survive challenge by active botulinum toxin F. This increase is typically mediated by an increased titre of antibodies to the toxin or an increased ability to produce antibodies to the toxin upon challenge with toxin. The term is not intended to indicate absolute protection against any amount of toxin.

The invention thus offers specific protection against a type F botulinum toxin, protection that has hitherto been unavailable.

In a particular embodiment the present invention provides a peptide or peptide conjugate comprising the amino acid sequence of the *C. botulinum* strain Langeland BoNT/F from amino acids 848 to 1278 (SEQ ID NO:1), but lacking the amino acid sequences of the L chain and H_N epitopes necessary for metalloprotease activity and toxin internalisation (found between amino acids 1 to 439 and 440 to 847, respectively); the peptide is capable of inducing an immune response that is protective against BoNT/F when administered to humans or other animals.

In a more particular embodiment the peptides of the invention consist of substantially only the sequence of amino acids from 848 to 1278 (SEQ ID NO:1) of the amino acid sequence of BoNT/F of the *Clostridium botulinum* strain Langeland, or of that sequence in the form of a fusion peptide with another amino acid sequence not being amino acids 1 to 847 of BoNT/F. The term `other amino acid sequence' will be understood by a person skilled in the art to include complete proteins as well as relatively short amino acid sequences as appropriate to the needs of the user. Optionally, the other amino acid sequence is a non–*C. botulinum*, antigenic protein which is included fused to the aforesaid sequence for the purpose of providing other immunity or labelling, or for modifying expression of the polypeptide in a host cell.

In another embodiment of the invention the polypeptide comprises a fragment or a derivative of a type F botulinum neurotoxin free of botulinum toxin activity and capable of induce protective immunity against type F toxin. The fragment is free of toxoid and

free of formaldehyde and has a length of less than 700 amino acids, preferably less than 500 amino acids.

In further specific embodiments of the invention the fragment is selected from:–

    (a)   amino acids 848–1278 of a type F botulinum toxin, (SEQ ID NO:1)

    (b)   amino acids 848–991 of a type F botulinum toxin, (SEQ ID NO:2)

    (c)   amino acids 992–1135 of a type F botulinum toxin, (SEQ ID NO:3) and

    (d)   amino acids 1136–1278 of a type F botulinum toxin (SEQ ID NO:4).

The invention also relates to a toxin derivative, being a synthetic polypeptide comprising a plurality of fragments of a type F botulinum toxin linked together in repeated sections. The derivative can comprise a dimer of the fragments specified above.

The first aspect of the invention also provides polypeptide compositions, free of botulinum toxin activity and capable of inducing protective immunity against botulinum toxin, which compositions are adapted so as to facilitate their processing. This is of advantage in the manufacture of vaccines as polypeptide must be separated out from a mixture of any components that are undesirable in an eventual vaccine. Such an adapted composition comprises.

    (1)   a polypeptide, free of botulinum toxin activity and capable of inducing protective immunity against a botulinum toxin; and

    (2)   a polypeptide adapted for purification of the composition.

Component (2) is adapted, for example, to facilitate purification of the composition from aqueous solution and optionally comprises an antibody, a binding region of an antibody, a polypeptide adapted to bind to an ion exchange column, a polypeptide adapted to bind to an affinity chromatography column or a purification ligand.

The composition preferably comprises or consists of a single polypeptide including

components (1) and (2), for example in the form of a fusion polypeptide.

In use of the compositions, extraction of the compositions from a mixture such as the supernatant from lysed cells expressing the composition is rendered a simple and fast process. It is particularly advantageous that in the composition, the vaccinating properties of component (1) are substantially retained, meaning that after purification of the composition it is used in a vaccine without the need for further modification, in particular without the need to remove component (2). As candidates for component (1) of the composition, all polypeptides previously described according to the first aspect of the invention are suitable. Further, fragments of tetanus toxin, free of toxin activity, are suitable.

A polypeptide according to a specific embodiment of the invention thus comprises a fusion protein of:-

(a)   amino acids 848 to 1278 (SEQ ID NO:1) of a type F botulinum neurotoxin, with

(b)   a purification moiety.

It is preferred that the purification moiety is adapted to bind to an affinity chromatography column. A typical purification moiety comprises from 50 to 500 amino acids. In a specific embodiment the fusion protein comprises maltose-binding protein as the purification moiety. This fusion protein is particularly suitable for purification using an affinity chromatography column and has been found to have useful vaccinating properties, as described below.

According to a second aspect the invention provides a vaccine against a botulinum toxin, comprising a polypeptide of the first aspect of the invention and a pharmaceutically acceptable carrier.

Suitable carriers are known to a person of skill in the for preparation of the vaccine. In an embodiment hereinafter described the carrier includes Freund's adjuvant. Another suitable carrier component is precipitated alum salts.

In a third aspect of the present invention there is provided a recombinant DNA encoding polypeptides of the invention. Such recombinant DNA is conveniently provided by PCR amplification of the DNA coding for the desired sequence, eg., BoNT/F$_{848-1278}$, using primers targeted at respective ends of the double stranded sequence. Optionally the template sequence used in PCR represents the natural clostridial gene. In a preferred embodiment of the invention, however, the sequence used is a synthetic sequence encoding the same amino acids as the natural clostridial protein but in which codon usage has been altered. It is preferred that the synthetic gene has a GC content of at least 40%, preferably at least 45% and most preferably at least 50%.

In the case of such a synthetic sequence, insertion into the chosen expression plasmid is achieved, in one embodiment of the invention, through the use of incorporated appropriate restriction endonuclease recognition sites positioned at the extremities of the DNA fragment during its construction.

By whatever means the recombinant DNA encoding the BoNT/F peptide is generated, it is ligated into a suitable expression vector at which stage genetic fusion to a desired fusion peptide encoding sequence occurs, if desired, and the resultant vector is introduced into a suitable cell line, eg., *E. coli* or a yeast such as *Pichia pastoris*. A cell line producing the desired product is selected through established procedures, eg., Western Blotting, or ELISA.

Fourth and fifth aspects of the invention provide respectively, a plasmid vector incorporating the DNA of the third aspect and a cell line comprising the plasmid and expressing the DNA.

The invention also provides a method for production of a toxin vaccine in which purification of active vaccinating agent is facilitated by its expression in combination with a polypeptide sequence adapted for purification. Accordingly, a sixth aspect of the invention provides a method for production of a toxin vaccine, said vaccine comprising a vaccinating polypeptide free of toxin activity and capable of inducing

protective immunity against a toxin, wherein the method comprises expressing in a host cell a DNA sequence coding for a fusion protein, said fusion protein comprising said vaccinating polypeptide and a purification moiety, obtaining an extract from the host cell comprising the fusion protein, and purifying therefrom the fusion protein.

In preferred embodiment of the sixth aspect of the invention there is provided a method of producing a vaccine containing a polypeptide of the first aspect of the invention, comprising the steps of:–

(a)   expressing in a host cell a DNA encoding a fusion protein, said protein being a fusion of (i) a fragment of a botulinum toxin, said fragment being free of toxin activity and capable of inducing protective immunity against botulinum toxin, and (ii) a purification moiety adapted to bind to an affinity chromatography column,

(b)   obtaining from said host cell an extract comprising the fusion protein, and

(c)   purifying the fusion protein using an affinity chromatography column.

In use of an embodiment of the invention the fusion protein is removed from the column by elution with a substrate. The method optionally includes cleaving the fusion protein and retaining the toxin fragment. The method has been used specifically with type F toxin but applies also to all other botulinum toxins and to tetanus toxin.

By this method the invention enables a preparation of botulinum toxin type F fragment free of contamination by other clostridial proteins, these latter frequently contaminating prior art preparations derived from cultures of *Clostridium* bacteria.

The fusion protein or toxin fragment obtained is typically in a substantially pure form and suitable for incorporation into a vaccine or other pharmaceutical composition in a few simple steps.

It should be noted that the creation of certain fusion proteins comprising the BoNT/F–derived peptide is useful in the initial isolation BoNT/F, following which

cleavage is optionally employed to purify the BoNT/F-related peptide. Where codons are added at the 5'-end of the BoNT/F-encoding DNA to aid in translation, these amino acids are optionally retained at the NH_2- terminal end of the final peptide, eg., those used to bring about secretion of the peptide or more simply the addition of an NH_2-terminal methionine to initiate translation.

A seventh aspect of the invention provides a method of making a pharmaceutical composition comprising:-

(a)   expressing in a host cell a DNA encoding a fusion protein, said protein being a fusion of (i) a botulinum toxin or a fragment thereof, free of toxin activity and capable of inducing protective immunity against botulinum toxin, and (ii)a purification moiety adapted to bind to an affinity chromatography column,

(b)   obtaining from said host cell an extract comprising the fusion protein,

(c)   purifying the fusion protein using an affinity chromatography column,

(d)   incorporating the purified fusion protein into a pharmaceutical composition.

The purification moiety typically comprises 50 to 500 amino acids, is water soluble and binds to an affinity chromatography column.

The inventors have found that a fusion protein retaining the purification moiety is of advantage in producing a vaccine against a type F botulinum toxin. Vaccinating activity is found in the fusion protein, so the purification protein does not need to be removed prior to vaccine manufacture, thus simplifying the manufacturing process. It is preferred that the purification protein is a globular, water soluble protein that binds to and can be purified using an affinity chromatography column. It is further preferred that the purification protein is highly immunogenic. Thus, a particularly preferred fusion protein comprises a fragment of a botulinum toxin free of toxin activity, an immunogenic region and a purification region adapted to bind to an affinity chromatography column.

The term immunogenic region is used above to describe a sequence of amino acids in a protein that is known to elicit stimulation of the immune system in humans or other animals. Examples of such an immunogenic region include keyhole limpet haemocyanin.

Further aspects of the invention provide a pharmaceutical containing the fusion protein, methods of vaccinating mammals using the vaccines and compositions of the invention and antisera raised against the polypeptides, vaccines and compositions of the invention.

There now follows description of specific embodiments of the invention, illustrated by drawings in which:–

Figure 1: shows the three major domains of a BoNT toxin. The numbers refer to the positions of the amino acids flanking these three domains in BoNT/F of *C.botulinum* strain Langeland;

Figure 2: shows a schematic representation of how synthetic gene blocks were assembled by PCR;

Figure 3: shows an example of a recombinant plasmid (pFHC206) made in which the synthetic DNA fragment in Figure 5 is inserted into the expression plasmid pMal–C2; and

Figure 4: shows antibody titres against BoNT/F obtained in mice immunised with MBP–BoNT/F$_{848-1278}$ recombinant protein.

SEQ ID NO:5 shows the nucleotide sequence of the region of the BoNT/F gene from *Clostridium botulinum* type F strain Langeland encoding the H$_C$ fragment;

SEQ ID NO:6 shows a synthetic DNA sequence encoding the BoNT/F H$_C$ fragment which uses codons which are used most frequently in highly expressed genes of E.

coli. The codon corresponding to BoNT/F Ser_{848} begins at nucleotide position 12. It is proceeded by a codon specifying a NH_2-terminal methionine codon and restriction sites for *Nde*I and *Bam*HI. The codon for Asn_{1278} begins at nucleotide position 1302, and is followed by a translational stop codon (nt 1305-1308) and a restriction site for *Xba*I;

EXAMPLES

<u>Generation of a synthetic DNA fragment encoding H_C of BoNT/F which makes use of codons utilised by highly expressed E. coli genes</u>

A synthetic sequence encoding BoNT/$F_{848-1278}$ was designed by reverse translation of the BoNT/F amino acid sequence using the REVTRANS programme of DNASTAR Inc (Madison, USA). The codon code used was the "strongly expressed E. coli backtranslation code" (SECOLI.RTC). To facilitate the construction, a number of changes were then made to introduce restriction enzyme recognition sites at strategic points along the length of the fragment, including unique flanking proximal sites for *Bam*HI and *Nde*I a distal flanking site for *Xba*I and internal sites for *Hpa*I, *Mlu*I and *Sp*I. The gene was then constructed from overlapping 100 mer oligonucleotides by a procedure essentially as described elsewhere [Sandhu *et al* (1992) Biotechniques 12:14-16].

Briefly, the gene was constructed as 4 individual blocks (A, B, C and D), each of approximately 350 bp in size. Each block was assembled from 4 x 100 mer alternating oligonucleotides which overlapped with each other by 20 nucleotides. These 4 oligonucleotides were used in a PCR to generate a composite c.350 bp double-stranded DNA fragment, which was subsequently amplified using 20 mer flanking primers. The amplified fragments of each block were then cloned directly into plasmid pCRII (Invitrogen Corp). The flanking primers of all 4 blocks were designed to include restriction enzyme sites which would allow their subsequent assembly into a contiguous fragment. Thus, block A was flanked by *Bam*HI (5') and *Hpa*I (3'), block B by *Hpa*I (5') and *Mlu*I (3'), block C by *Mlu*I (5') and *Sp*1L (3'), and block D by *Sp*1I

(5') and *Xba*I (3'). Each block was, therefore, released from their respective pCRII–derived recombinant plasmid by cleavage with the appropriate enzyme and the isolated fragments ligated to pMTL23 [Chambers *et al* (1988). Gene 68:139–149] plasmid DNA which had been cleaved with *Bam*HI and *Xba*I. A clone was then selected in which all 4 blocks had been inserted in the expected order. This was confirmed by nucleotide sequencing using routine methods [Maniatis *et al*. (1989). Molecular Cloning a Laboratory Manual. Cold Spring Harbor Laboratory Press], and the plasmid obtained designed pFHC23.

Generation of a H_C peptide (848 to 1278) of BoNT/F of C. botulinum strain Langeland

A candidate vaccine against the BoNT/F of *C. botulinum* was produced by expressing the fragment of the synthetic gene encoding the H_C fragment, amino acids 848 to 1278. This DNA fragment was isolated from plasmid pFHC23 as an approximately 1.3 kb *Bam*HI–*Xho*I restriction fragment and inserted between the unique *Bam*HI and *Sal*I sites of pUC9 [Vieira and Messing (1982). Gene 19: 259–268], generating the plasmid pFHC29. The insert was then reisolated from pFHC29 as an *Eco*RI–*Xba*I fragment and inserted between the equivalent sites of the commercially available expression vector pMal–c2 (New England Biolabs), to yield the final plasmid pFHC206. The resultant plasmid expressed BoNT/F$_{848-1278}$ as a fusion protein with the vector encoded maltose binding protein (MBP).

Fusion protein product (MBP–BoNT/F$_{848-1278}$) was prepared from the cell line containing pFHC206 in the following manner. *E. coli* containing pFHC206 was cultivated in 1 litre of media (M9, supplemented with 0.8M sorbitol, 0.5% casamino acids, 50 μg/ml ampicillin), shaking (200 rpm) at 37^0C until an OD$_{600}$ of 1.0 was achieved. At this point IPTG was added at a final concentration of 1 mM and shaking continued at 27^0C for a further 4 hour. Cells were harvested by centrifugation (5000 x *g*) and resuspended in 20 ml of lysis buffer (Protein Fusion and Purification System, New England Biolabs) and cells disrupted by sonication. Lysate was applied to a GPC column containing 180 ml of Sephacel S100, and the protein in the void fraction collected. MBP–BoNT/F H$_{848-1278}$ fusion protein in this fraction was then allowed

to adsorb at room temperature to a 4–6 ml volume of Amylose resin (New England Laboratories) over a 3 hour period with gentle shaking (10 rpm). Recombinant fusion protein was then eluted in buffer (0.01M Tris HCl, pH 7.0) containing 5 mM maltose. Eluted protein was concentrated using an Amicon PM30 membrane filter.

Toxicity of candidate vaccine

The toxicity of the candidate vaccine fusion peptide was determined by intraperitoneal inoculation of 25 μg amounts of the total recombinant MBP-BoNT/F$_{848-1278}$ protein into groups of 4 mice. The candidate vaccine was well tolerated and mice showed no signs of acute or chronic toxicity up to 2 weeks post inoculation.

Antibody responses to candidate vaccines

The candidate vaccine was administered to groups of 4 mice by intraperitoneal inoculation in complete Freund's adjuvant, and a booster inoculation given on 3 further occasions at two week intervals. Antibody response against purified *C. botulinum* strain Langeland BoNT/F was monitored by ELISA (Fig 4).

Protection against toxin challenge

Animals which were immunised with MBP-BoNT/F$_{848-1278}$ fusion protein were subjected to an intraperitoneal challenge with various doses of purified *C. botulinum* strain Langeland BoNT/F. At doses of 12 LD$_{50}$ and above, all the control, unimmunised mice succumbed within 24 hour. All immunised groups of mice survived challenges of up to 2.4 X 10^4 LD$_{50}$. One of the immunised mice which had survived an initial challenge of 1.8, LD$_{50}$ was subsequently shown to be immune to a further challenge of 10^6 LD$_{50}$.

TABLE 1: Protection against challenge with *C. botulinum* strain Langeland BoNT/F afforded by the MBP-BoNT/F$_{848-1278}$ fusion protein vaccine. A total of 4 X 25 μg intraperitoneal doses of antigen mixed with adjuvant were given to groups of 4 mice

- 14 -

at 14 day intervals. After 50 days mice were subjected to intraperitoneal challenges of varying levels of purified BoNT/F, (isolated from *C. botulinum* strain Langeland), and deaths recorded up to 4 days.

Challenge Dose (LD$_{50}$)	Mortality/Total Animals	
	Control Animals	Immunised Animals
2.4×10^4	4/4	0/4
3.6×10^3	4/4	0/4
5.4×10^2	4/4	0/4
81	4/4	0/4
12	4/4	0/4
1.8	2/4	0/4[a]

[a] = one of the surviving individuals from this group was subsequently shown to be protected against a BoNT/F challenge equivalent to 10^6 LD$_{50}$.

This invention provides a fragment (such as amino acids 848-1278) of BoNT/F isolated from *C. botulinum* strain Langeland for use as a vaccine. The fragment retains its immunogenic properties while still fused with MBP, dispensing with the need for an additional purification step. The recombinant fusion protein appears to be non-toxic in mice at doses up to 25µg. Repeated doses produced a significant antibody response which protects animals against BoNT/F challenge. As a vaccine it offers several advantages over neurotoxin toxoided by formaldehyde treatment. Most notably, it may be prepared more easily and, due to the absence of active toxin, at a lower level of containment. The absence of other contaminating *C.botulinum* proteins and partially toxoided materials also make it inherently safer for vaccine application and less reactogenic.

SEQUENCE LISTING.

(1) GENERAL INFORMATION:

    (i) APPLICANT:
        (A) NAME: Microbiological Research Authority
        (B) STREET: Centre for Applied Microbiology and Research,
               Porton Down
        (C) CITY: Salisbury
        (D) STATE: Wilshire
        (E) COUNTRY: UK
        (F) POSTAL CODE (ZIP): SP4 OJG

        (A) NAME: Nigel Peter Minton
        (B) STREET: 27 Moberly Road
        (C) CITY: Salisbury
        (D) STATE: Wiltshire
        (E) COUNTRY: UK
        (F) POSTAL CODE (ZIP): SP1 3BZ

        (A) NAME: Michael J Elmore
        (B) STREET: 8 St Mary's Court, Eastrop Lane
        (C) CITY: Basingstoke
        (D) STATE: Hants
        (E) COUNTRY: UK
        (F) POSTAL CODE (ZIP): RG21 4AT

        (A) NAME: Margaret Lamble Mauchline
        (B) STREET: Three Dormers, Bakers Hill, Semley
        (C) CITY: Shaftesbury
        (D) STATE: Dorset
        (E) COUNTRY: UK
        (F) POSTAL CODE (ZIP): SP7 9BQ

        (A) NAME: Vladimir Artymovich Pasechnik
        (B) STREET: 1 Copper Beech Close
        (C) CITY: Shrewton
        (D) STATE: Wiltshire
        (E) COUNTRY: UK
        (F) POSTAL CODE (ZIP): SP4 4HU

    (ii) TITLE OF INVENTION: BOTULINUM TOXIN VACCINE AND ITS MANUFACTURE

   (iii) NUMBER OF SEQUENCES: 6

    (iv) COMPUTER READABLE FORM:
        (A) MEDIUM TYPE: Floppy disk
        (B) COMPUTER: IBM PC compatible
        (C) OPERATING SYSTEM: PC-DOS/MS-DOS
        (D) SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPO)

(2) INFORMATION FOR SEQ ID NO: 1:

(i) SEQUENCE CHARACTERISTICS:
    (A) LENGTH: 431 amino acids
    (B) TYPE: amino acid
    (C) STRANDEDNESS: single
    (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1:

```
Ser Tyr Thr Asn Asp Lys Ile Leu Ile Leu Tyr Phe Asn Lys Leu Tyr
1               5               10              15

Lys Lys Ile Lys Asp Asn Ser Ile Leu Asp Met Arg Tyr Glu Asn Asn
        20              25              30

Lys Phe Ile Asp Ile Ser Gly Tyr Gly Ser Asn Ile Ser Ile Asn Gly
        35              40              45

Asp Val Tyr Ile Tyr Ser Thr Asn Arg Asn Gln Phe Gly Ile Tyr Ser
    50              55              60

Ser Lys Pro Ser Glu Val Asn Ile Ala Gln Asn Asn Asp Ile Ile Tyr
65              70              75              80

Asn Gly Arg Tyr Gln Asn Phe Ser Ile Ser Phe Trp Val Arg Ile Pro
                85              90              95

Lys Tyr Phe Asn Lys Val Asn Leu Asn Asn Glu Tyr Thr Ile Ile Asp
            100             105             110

Cys Ile Arg Asn Asn Asn Ser Gly Trp Lys Ile Ser Leu Asn Tyr Asn
        115             120             125

Lys Ile Ile Trp Thr Leu Gln Asp Thr Ala Gly Asn Asn Gln Lys Leu
    130             135             140

Val Phe Asn Tyr Thr Gln Met Ile Ser Ile Ser Asp Tyr Ile Asn Lys
145             150             155             160

Trp Ile Phe Val Thr Ile Thr Asn Asn Arg Leu Gly Asn Ser Arg Ile
            165             170             175

Tyr Ile Asn Gly Asn Leu Ile Asp Glu Lys Ser Ile Ser Asn Leu Gly
        180             185             190

Asp Ile His Val Ser Asp Asn Ile Leu Phe Lys Ile Val Gly Cys Asn
        195             200             205

Asp Thr Arg Tyr Val Gly Ile Arg Tyr Phe Lys Val Phe Asp Thr Glu
    210             215             220
```

Leu Gly Lys Thr Glu Ile Glu Thr Leu Tyr Ser Asp Glu Pro Asp Pro
225   230   235   240

Ser Ile Leu Lys Asp Phe Trp Gly Asn Tyr Leu Leu Tyr Asn Lys Arg
            245   250   255

Tyr Tyr Leu Leu Asn Leu Leu Arg Thr Asp Lys Ser Ile Thr Gln Asn
        260   265   270

Ser Asn Phe Leu Asn Ile Asn Gln Gln Arg Gly Val Tyr Gln Lys Pro
        275   280   285

Asn Ile Phe Ser Asn Thr Arg Leu Tyr Thr Gly Val Glu Val Ile Ile
    290   295   300

Arg Lys Asn Gly Ser Thr Asp Ile Ser Asn Thr Asp Asn Phe Val Arg
305   310   315   320

Lys Asn Asp Leu Ala Tyr Ile Asn Val Val Asp Arg Asp Val Glu Tyr
            325   330   335

Arg Leu Tyr Ala Asp Ile Ser Ile Ala Lys Pro Glu Lys Ile Ile Lys
        340   345   350

Leu Ile Arg Thr Ser Asn Ser Asn Asn Ser Leu Gly Gln Ile Ile Val
        355   360   365

Met Asp Ser Ile Gly Asn Asn Cys Thr Met Asn Phe Gln Asn Asn Asn
    370   375   380

Gly Gly Asn Ile Gly Leu Leu Gly Phe His Ser Asn Asn Leu Val Ala
385   390   395   400

Ser Ser Trp Tyr Tyr Asn Asn Ile Arg Lys Asn Thr Ser Ser Asn Gly
            405   410   415

Cys Phe Trp Ser Phe Ile Ser Lys Glu His Gly Trp Gln Glu Asn
            420   425   430

(2) INFORMATION FOR SEQ ID NO: 2:

    (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 144 amino acids
        (B) TYPE: amino acid
        (C) STRANDEDNESS: single
        (D) TOPOLOGY: linear

    (ii) MOLECULE TYPE: peptide

    (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2:

    Ser Tyr Thr Asn Asp Lys Ile Leu Ile Leu Tyr Phe Asn Lys Leu Tyr
    1   5   10   15

- 18 -

```
Lys Lys Ile Lys Asp Asn Ser Ile Leu Asp Met Arg Tyr Glu Asn Asn
            20              25              30

Lys Phe Ile Asp Ile Ser Gly Tyr Gly Ser Asn Ile Ser Ile Asn Gly
            35              40              45

Asp Val Tyr Ile Tyr Ser Thr Asn Arg Asn Gln Phe Gly Ile Tyr Ser
            50              55              60

Ser Lys Pro Ser Glu Val Asn Ile Ala Gln Asn Asn Asp Ile Ile Tyr
65              70              75              80

Asn Gly Arg Tyr Gln Asn Phe Ser Ile Ser Phe Trp Val Arg Ile Pro
                85              90              95

Lys Tyr Phe Asn Lys Val Asn Leu Asn Asn Glu Tyr Thr Ile Ile Asp
            100             105             110

Cys Ile Arg Asn Asn Asn Ser Gly Trp Lys Ile Ser Leu Asn Tyr Asn
            115             120             125

Lys Ile Ile Trp Thr Leu Gln Asp Thr Ala Gly Asn Asn Gln Lys Leu
    130             135             140
```

(2) INFORMATION FOR SEQ ID NO: 3:

    (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 144 amino acids
        (B) TYPE: amino acid
        (C) STRANDEDNESS: single
        (D) TOPOLOGY: linear

    (ii) MOLECULE TYPE: peptide

    (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 3:

```
Val Phe Asn Tyr Thr Gln Met Ile Ser Ile Ser Asp Tyr Ile Asn Lys
1               5               10              15

Trp Ile Phe Val Thr Ile Thr Asn Asn Arg Leu Gly Asn Ser Arg Ile
            20              25              30

Tyr Ile Asn Gly Asn Leu Ile Asp Glu Lys Ser Ile Ser Asn Leu Gly
            35              40              45

Asp Ile His Val Ser Asp Asn Ile Leu Phe Lys Ile Val Gly Cys Asn
            50              55              60

Asp Thr Arg Tyr Val Gly Ile Arg Tyr Phe Lys Val Phe Asp Thr Glu
65              70              75              80
```

```
Leu Gly Lys Thr Glu Ile Glu Thr Leu Tyr Ser Asp Glu Pro Asp Pro
            85                90                95

Ser Ile Leu Lys Asp Phe Trp Gly Asn Tyr Leu Leu Tyr Asn Lys Arg
            100               105           110

Tyr Tyr Leu Leu Asn Leu Leu Arg Thr Asp Lys Ser Ile Thr Gln Asn
        115               120           125

Ser Asn Phe Leu Asn Ile Asn Gln Gln Arg Gly Val Tyr Gln Lys Pro
    130               135           140
```

(2) INFORMATION FOR SEQ ID NO: 4:

   (i) SEQUENCE CHARACTERISTICS:
       (A) LENGTH: 143 amino acids
       (B) TYPE: amino acid
       (C) STRANDEDNESS: single
       (D) TOPOLOGY: linear

   (ii) MOLECULE TYPE: peptide

   (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 4:

```
Asn Ile Phe Ser Asn Thr Arg Leu Tyr Thr Gly Val Glu Val Ile Ile
1               5               10              15

Arg Lys Asn Gly Ser Thr Asp Ile Ser Asn Thr Asp Asn Phe Val Arg
            20              25              30

Lys Asn Asp Leu Ala Tyr Ile Asn Val Val Asp Arg Asp Val Glu Tyr
        35              40              45

Arg Leu Tyr Ala Asp Ile Ser Ile Ala Lys Pro Glu Lys Ile Ile Lys
    50              55              60

Leu Ile Arg Thr Ser Asn Ser Asn Asn Ser Leu Gly Gln Ile Ile Val
65              70              75              80

Met Asp Ser Ile Gly Asn Asn Cys Thr Met Asn Phe Gln Asn Asn Asn
            85              90              95

Gly Gly Asn Ile Gly Leu Leu Gly Phe His Ser Asn Asn Leu Val Ala
            100             105           110

Ser Ser Trp Tyr Tyr Asn Asn Ile Arg Lys Asn Thr Ser Ser Asn Gly
        115             120           125

Cys Phe Trp Ser Phe Ile Ser Lys Glu His Gly Trp Gln Glu Asn
    130             135           140
```

(2) INFORMATION FOR SEQ ID NO: 5:

    (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 1293 base pairs
        (B) TYPE: nucleic acid
        (C) STRANDEDNESS: double
        (D) TOPOLOGY: linear

    (ii) MOLECULE TYPE: DNA (genomic)

    (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 5:

```
TCATATACTA ATGATAAAAT TCTAATTTTA TATTTTAATA AATTATATAA AAAAATTAAA      60

GATAACTCTA TTTTAGATAT GCGATATGAA AATAATAAAT TTATAGATAT CTCTGGATAT     120

GGTTCAAATA TAAGCATTAA TGGAGATGTA TATATTTATT CAACAAATAG AAATCAATTT     180

GGAATATATA GTAGTAAGCC TAGTGAAGTT AATATAGCTC AAAATAATGA TATTATATAC     240

AATGGTAGAT ATCAAAATTT TAGTATTAGT TTCTGGGTAA GGATTCCTAA ATACTTCAAT     300

AAAGTGAATC TTAATAATGA ATATACTATA ATAGATTGTA TAAGGAATAA TAATTCAGGA     360

TGGAAAATAT CACTTAATTA TAATAAAATA ATTTGGACTT TACAAGATAC TGCTGGAAAT     420

AATCAAAAAC TAGTTTTTAA TTATACACAA ATGATTAGTA TATCTGATTA TATAAATAAA     480

TGGATTTTTG TAACTATTAC TAATAATAGA TTAGGCAATT CTAGAATTTA CATCAATGGA     540

AATTTAATAG ATGAAAAATC AATTTCGAAT TTAGGTGATA TTCATGTTAG TGATAATATA     600

TTATTTAAAA TTGTTGGTTG TAATGATACA AGATATGTTG GTATAAGATA TTTTAAAGTT     660

TTTGATACGG AATTAGGTAA AACAGAAATT GAGACTTTAT ATAGTGATGA GCCAGATCCA     720

AGTATCTTAA AAGACTTTTG GGGAAATTAT TTGTTATATA ATAAAAGATA TTATTTATTG     780

AATTTACTAA GAACAGATAA GTCTATTACT CAGAATTCAA ACTTTCTAAA TATTAATCAA     840

CAAAGAGGTG TTTATCAGAA ACCAAATATT TTTTCCAACA CTAGATTATA TACAGGAGTA     900

GAAGTTATTA TAAGAAAAAA TGGATCTACA GATATATCTA ATACAGATAA TTTTGTTAGA     960

AAAAATGATC TGGCATATAT TAATGTAGTA GATCGTGATG TAGAATATCG GCTATATGCT    1020

GATATATCAA TTGCAAAACC AGAGAAAATA ATAAAATTAA TAAGAACATC TAATTCAAAC    1080

AATAGCTTAG GTCAAATTAT AGTTATGGAT TCAATAGGAA ATAATTGCAC AATGAATTTT    1140

CAAAACAATA ATGGGGGCAA TATAGGATTA CTAGGTTTTC ATTCAAATAA TTTGGTTGCT    1200
```

AGTAGTTGGT ATTATAACAA TATACGAAAA AATACTAGCA GTAATGGATG CTTTTGGAGT   1260

TTTATTTCTA AAGAGCATGG ATGGCAAGAA AAC   1293

(2) INFORMATION FOR SEQ ID NO: 6:

    (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 1313 base pairs
        (B) TYPE: nucleic acid
        (C) STRANDEDNESS: double
        (D) TOPOLOGY: linear

    (ii) MOLECULE TYPE: DNA (genomic)

    (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 6:

GGATCCATAT GTCTTACACT AACGACAAAA TCCTGATCCT GTACTTCAAC AAAACTGTACA   60

AAAAAATCAA AGACAACTCT ATCCTGGACA TGCGTTACGA AAACAACAAA TTCATCGACA   120

TCTCTGGCTA TGGTTCTAAC ATCTCTATCA ACGGTGACGT CTACATCTAC TCTACTAACC   180

GCAACCAGTT CGGTATCTAC TCTTCTAAAC CGTCTGAAGT AAACATCGCT CAGAACAACG   240

ACATCATCTA CAACGGTCGT TACCAGAACT TCTCTATCTC TTTCTGGGTT CGTATCCCGA   300

AATACTTCAA CAAAGTTAAC CTGAACAACG AATACACTAT CATCGACTGC ATCCGTAACA   360

ACAACTCTGG TTGGAAAATC TCTCTGAACT ACAACAAAAT CATCTGGACT CTGCAGGACA   420

CTGCTGGTAA CAACCAGAAA CTGGTTTTCA ACTACACTCA GATGATCTCT ATCTCTGACT   480

ACATTAATAA ATGGATCTTC GTTACTATCA CTAACAACCG TCTGGGTAAC TCTCGTATCT   540

ACATCAACGG TAACCTGATC GATGAAAAAT CTATCTCTAA CCTGGGTGAC ATCCACGTTT   600

CTGACAACAT CCTGTTCAAA ATCGTTGGTT GCAACGACAC GCGTTACGTT GGTATCCGTT   660

ACTTCAAAGT TTTCGACACT GAACTGGGTA AAACTGAAAT CGAAACTCTG TACTCTGACG   720

AACCGGACCC GTCTATCCTG AAAGACTTCT GGGGTAACTA CCTGCTGTAC AACAAACGTT   780

ACTACCTGCT GAACCTGCTC CGGACTGACA AATCTATCAC TCAGAACTCT AACTTCCTGA   840

ACATCAACCA GCAGCGTGGT GTTTATCAGA AACCTAATAT CTTCTCTAAC ACTCGTCTGT   900

ACACTGGTGT TGAAGTTATC ATCCGTAAAA ACGGTTCTAC TGACATCTCT AACACTGACA   960

ACTTCGTACG TAAAAACGAC CTGGCTTACA·TCAACGTTGT TGACCGTGAC GTTGAATACC   1020

GTCTGTACGC TGACATCTCT ATCGCTAAAC CGGAAAAAAT CATCAAACTG ATCCGTACTT   1080

```
CTAACTCTAA CAACTCTCTG GGTCAGATCA TCGTTATGGA CTCGATCGGT AACAACTGCA      1140

CTATGAACTT CCAGAACAAC AACGGTGGTA ACATCGGTCT GCTGGGTTTC CACTCTAACA      1200

ACCTGGTTGC TTCTTCTTGG TACTACAACA ACATCCGTAA AAACACTTCT TCTAACGGTT      1260

GCTTCTGGTC TTTCATCTCT AAAGAACACG GTTGGCAGGA AAACTAATCT AGA             1313
```

CLAIMS

1.   A polypeptide free of botulinum toxin activity which induces protective immunity
to a type F botulinum toxin.

2.   A polypeptide characterized in that it:–
    (a)   is free of botulinum toxin activity, and
    (b)   is capable of eliciting, in a mammal, an immunological response that is
    protective against type F botulinum toxin.

3.   A polypeptide according to Claim 1 or 2 comprising a fragment or a derivative
of a heavy chain of a type F botulinum neurotoxin.

4.   A polypeptide according to Claim 3 wherein said fragment or said derivative is
up to 600 amino acids long.

5.   A polypeptide according to Claims 3 or 4 wherein said fragment is selected
from:–
    (a)   amino acids 848–1278 of a type F botulinum toxin,
    (b)   amino acids 848–991 of a type F botulinum toxin,
    (c)   amino acids 992–1135 of a type F botulinum toxin, and
    (d)   amino acids 1136–1278 of a type F botulinum toxin.

6.   A polypeptide according to Claims 3 or 4 wherein said derivative comprises a
dimer of the fragment according to any of (a)–(d) of Claim 5.

7.   A polypeptide composition for use in manufacture of a vaccine, said
composition comprising:–

    (1)   a polypeptide free of toxin activity and capable of inducing, in a mammal,
    protective immunity against a botulinum toxin or a tetanus toxin; and

(2)   a polypeptide adapted to facilitate or enhance purification of the composition.

8.   A polypeptide composition according to Claim 7 wherein the composition comprises a fusion protein.

9.   A polypeptide composition according to Claim 7 or 8 comprising:-

(1)   a polypeptide according to any of Claims 1-6; and

(2)   a polypeptide adapted to bind to a chromatography column.

10.   A polypeptide composition according to any of Claims 7-9 comprising a polypeptide adapted to bind to an affinity chromatography column.

11.   A polypeptide according to Claim 8 comprising a fusion protein of:-

(a)   amino acids 848 to 1278 of a type F botulinum neurotoxin, with

(b)   a purification moiety.

12.   A vaccine comprising a pharmaceutically acceptable carrier and a polypeptide according to any of Claims 1-6 or a polypeptide composition according to any of Claims 7-11.

13.   A recombinant DNA encoding a polypeptide according to any of Claims 1-6 or a polypeptide composition according to any of Claims 7-11.

14.   A method of producing a polypeptide according to any of Claims 1-6 or a polypeptide composition according to any of Claims 7-11 comprising the steps of:-

(a)   expressing in a host cell a DNA encoding a fusion protein, said protein being a fusion of (i) a fragment or derivative of a type F botulinum toxin, and (ii) a moiety adapted to bind to an affinity chromatography column,

(b)   obtaining from said host cell an extract comprising the fusion protein, and

(c)   purifying the fusion protein using an affinity chromatography column.

15.   A method according to Claim 14 wherein the fusion protein is removed from the column by elution with a substrate.

16.   A method according to Claim 14 or 15 further comprising cleaving the fusion protein and retaining the toxin fragment.

17.   A method of making a pharmaceutical composition comprising:–

(a)   expressing in a host cell a DNA encoding a fusion protein, said protein being a fusion of (i) a polypeptide free of toxin activity and capable of inducing protective immunity against a botulinum toxin or a tetanus toxin, and (ii) a purification moiety adapted to bind to a chromatography column,

(b)   obtaining from said host cell an extract comprising the fusion protein,

(c)   purifying the fusion protein using chromatography column,

(d)   incorporating the purified fusion protein into a pharmaceutical composition.

18.   A method according to Claim 17 wherein said purification moiety binds to an affinity chromatography column.

19.   A pharmaceutical composition comprising:–

(a)   a fusion protein, said protein being a fusion of (i) a polypeptide fee of toxin activity and capable of inducing protective immunity against a botulinum toxin or a tetanus toxin, and (ii) a polypeptide adapted to bind to a chromatography column; and

(b)   a pharmaceutically acceptable carrier.

20.   A pharmaceutical composition according to Claim 19 wherein said fusion
        protein comprises a polypeptide according to any of Claims 1–6.

21.   A pharmaceutical composition according to Claim 19 or 20 wherein the fusion
        protein comprises a polypeptide adapted to bind to an affinity chromatography
        column.

22.   A method of vaccinating a mammal against a botulinum toxin, comprising
        administering to said mammal a vaccine according to Claim 12.

23.   A method of vaccinating a mammal against a botulinum toxin, comprising
        administering to said mammal a pharmaceutical composition according to any
        of Claims 19–21.
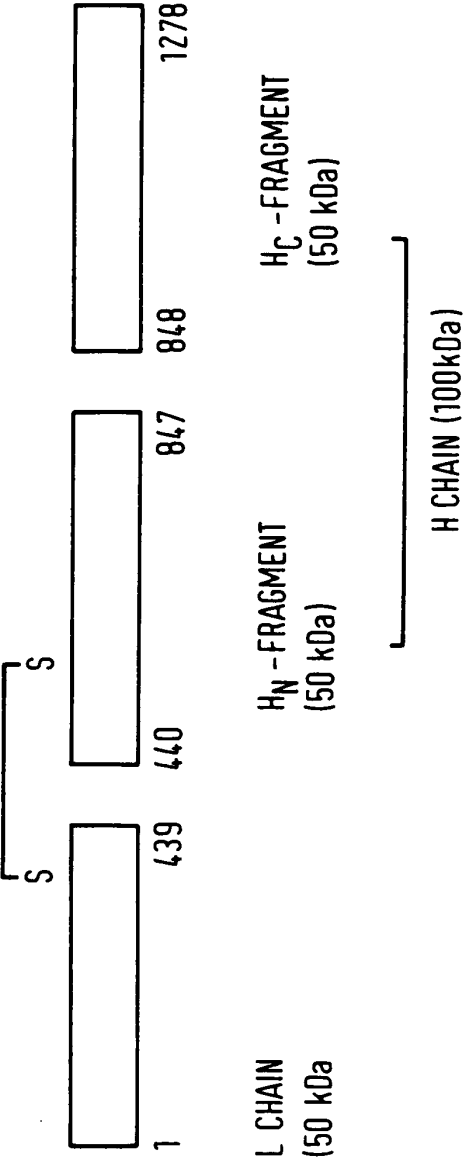
FIG. 1
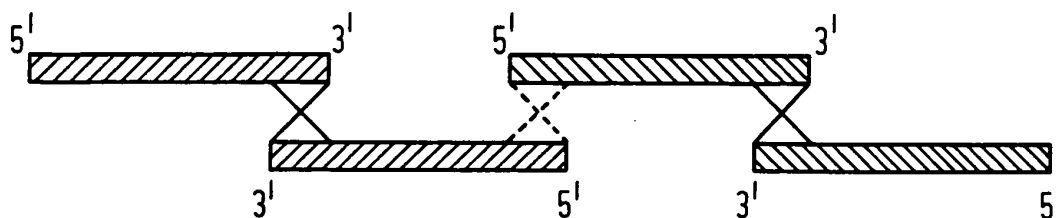
FIG. 2

pMAL™-c2,-p2:

POLYLINKER

malE   lacZα

rrnB
TERMINATOR

P_{tac}

pMAL-c2
pMAL-p2

AMP'

lacI^q

+
M13 ori
-

pBR322 ori

pMAL™-c2,-p2: POLYINKER

Xmnl   EcoRI BamHI Xbal   Sall   Pstl   HindIII

malE..ATC GAG GGA AGG ATT TCA GAA TTC GGA TCC TCT AGA GTC GAC CTG CAG GCA AGC TTG...lacZα
      ile glu gly arg ile

FACTOR Xa
CLEAVAGE SITE

BoNT/F-H_c

EcoRI   Hpal   Mlul   Spll   Xbal

0   0.2   0.4   0.6   0.8   1.0   1.2   1.3-kb
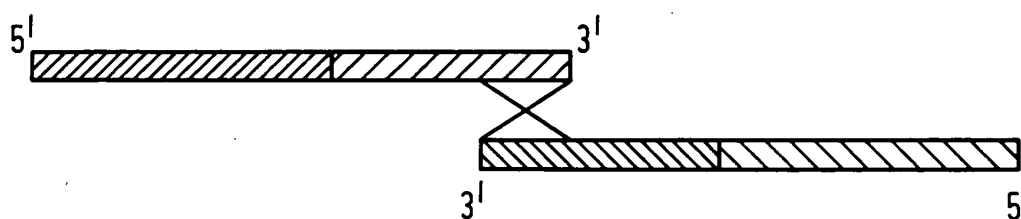
FIG. 3

4/4

FIG. 4

ANTIGENICITY OF SERUM AFTER IMMUNISATION OF MICE WITH
MBP-BoNT/F (848-1278) RECOMBINANT PROTEIN
ANTIGEN: BoNT/F, 13ng/WELL
SERA S-2 (1) AND S-3 (2) WERE AFTER SECOND AND THIRD BOOSTS
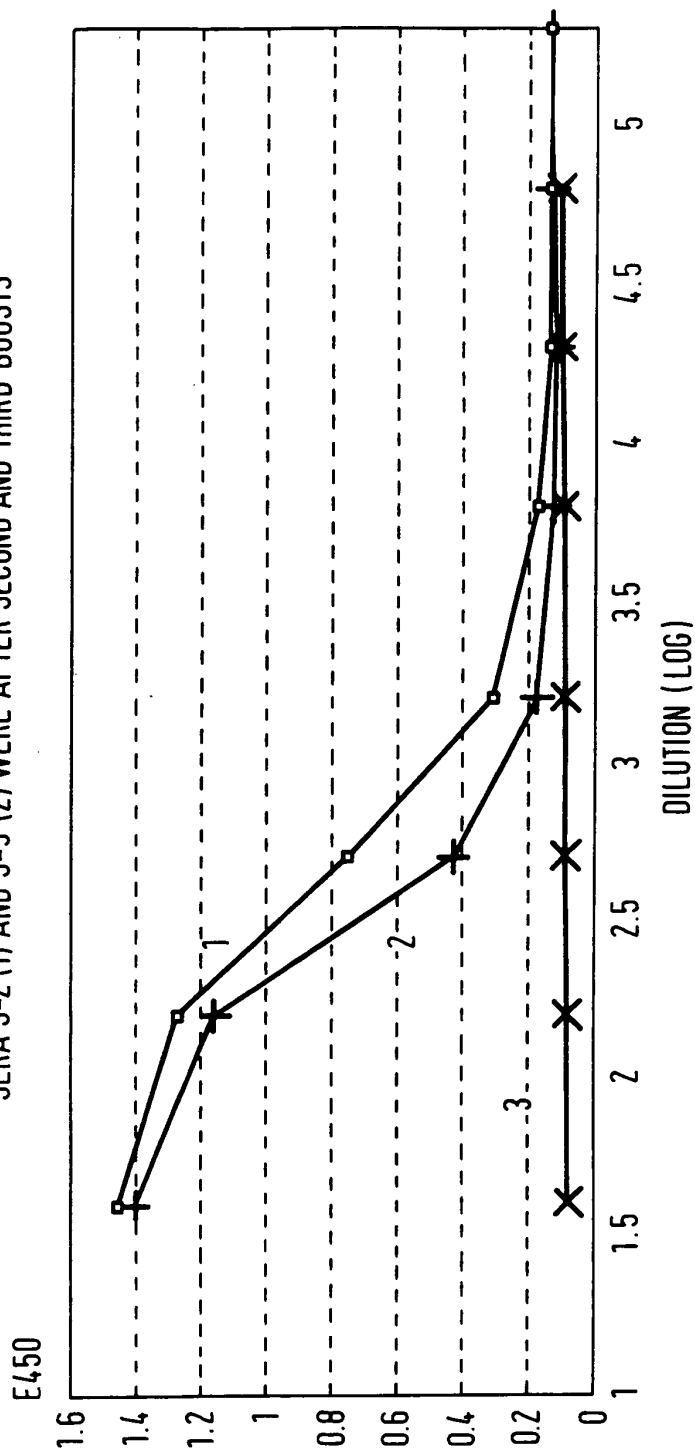
SERUM S-2, S-3 AS WELL AS NON-IMMUNE SERUM WERE FIRST DILUTED 1:50 AND 1:3
AT EACH NEXT STEP

(3): NON-IMMUNE SERA

INTERNATIONAL SEARCH REPORT

A. CLASSIFICATION OF SUBJECT MATTER
IPC 6   C12N15/31   C12N15/62   C07K14/33   A61K39/08

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)
IPC 6   C07K

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

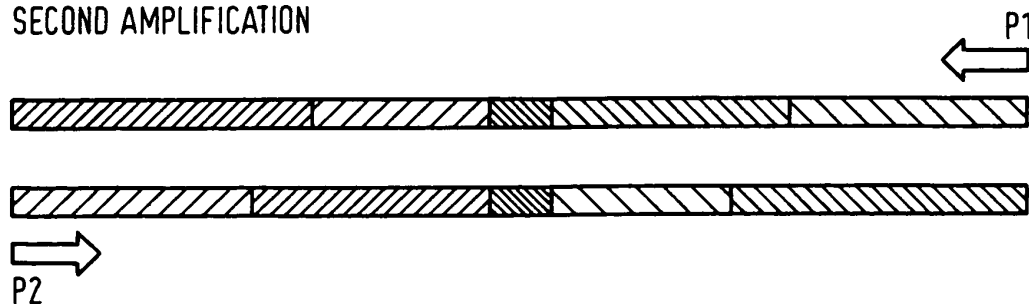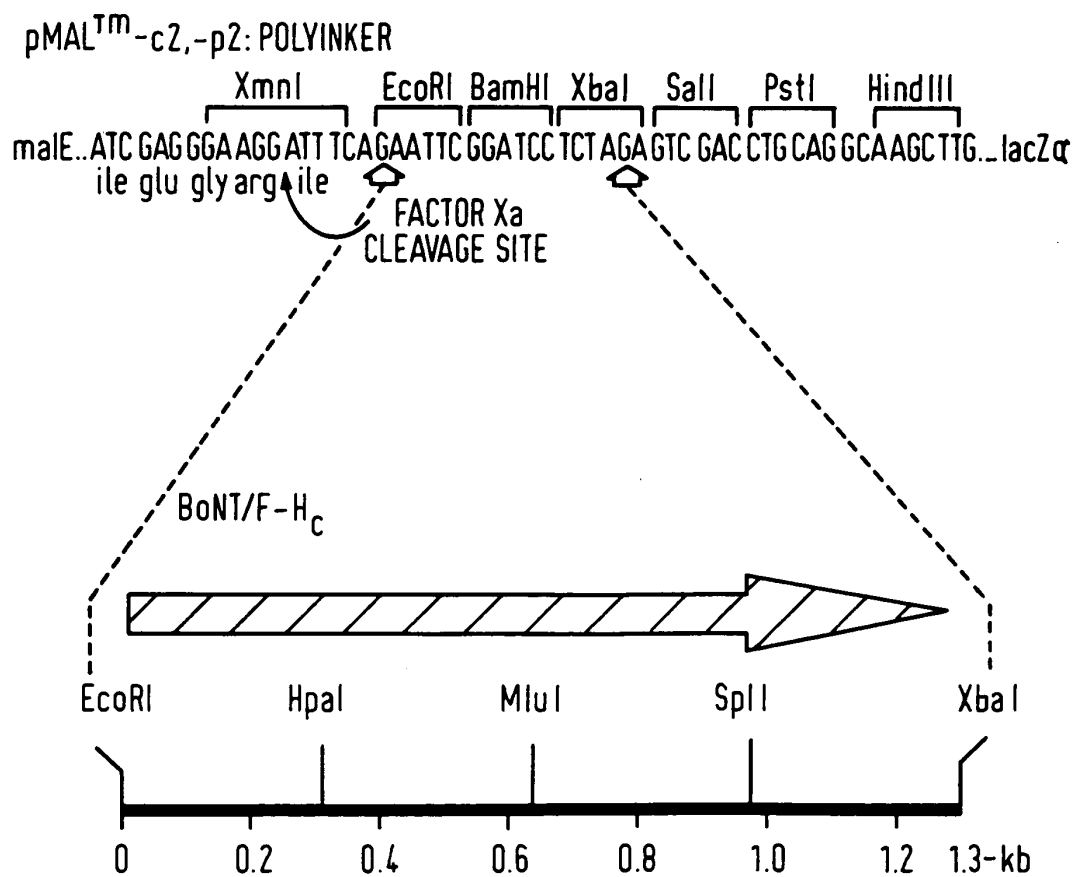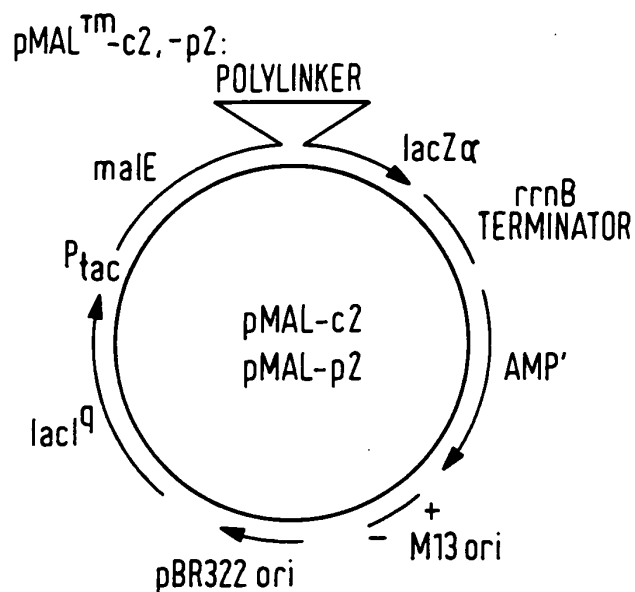Electronic data base consulted during the international search (name of data base and, where practical, search terms used)

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	DATABASE MEDLINE FILE SERVER STN KARLSRUHE ABSTRACT 77064466, HATHEWAY: "TOXOID OF CLOSTRIDIUM BOTULINUM TYPE F: PURIFICATION AND IMMUNOGENICITY STUDIES" XP002015940 & APPLIED AND ENVIROMENTAL MICROBIOLOGY, (1976 FEB) 31 (2) 234-42. see abstract ---	1,2,12, 22
X	ABSTRACTS OF THE 95TH GENERAL MEETING OF THE AMERICAN SOCIETY FOR MICROBIOLOGY, 21 - 25 May 1995, WASHINGTON D.C,USA, page 289 XP002015937 MONTGOMERY ET AL: "EVALUATION OF BOTULINUM TYPE F VACCINE BY ELISA" see abstract E-49 ---	1,2,12, 22

-/--

X	Further documents are listed in the continuation of box C.		X	Patent family members are listed in annex.

* Special categories of cited documents :

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier document but published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.

"&" document member of the same patent family

Date of the actual completion of the international search	Date of mailing of the international search report
16 October 1996	0 4. 11. 96

Name and mailing address of the ISA	Authorized officer
European Patent Office, P.B. 5818 Patentlaan 2 NL - 2280 HV Rijswijk Tel. (+ 31-70) 340-2040, Tx. 31 651 epo nl, Fax (+ 31-70) 340-3016	Sitch, W

Form PCT/ISA/210 (second sheet) (July 1992)

page 1 of 3

C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	BIOCHEMISTRY, vol. 33, 1994, pages 7014-7020, XP002015938 LI ET AL: "A SINGLE MUTATION IN THE RECOMBINANT LIGHT CHAIN OF TETANUS TOXIN ABOLISHES ITS PROTEOLYTIC ACTIVITY AND REMOVES THE TOXICITY SEEN AFTER RECONSTITUTION WITH NATIVE HEAVY CHAIN" see the whole document ---	7,8,10, 12,13, 17-19,21
X	WO,A,94 03615 (MEDEVA HOLDINGS B V ;KHAN MOHAMMED ANJAM (GB); HORMAECHE CARLOS ES) 17 February 1994 see page 10, paragraph 2 - page 11, paragraph 2 ---	7,8,10, 12,13, 17-19,21
X	DATABASE CHEMICAL ABSTRACTS FILE SERVER STN KARLSRUHE ABSTRACT 123:219673, MINTON: "PHYSICAL CHARACTERIZATION OF CLOSTRIDIUM BOTULINUM NEUROTOXIN GENES" XP002015941 & REPORT (1993),ORDER NO.AD-A272939 ... see abstract ---	7,8,10, 12,13, 17-19, 21-23
A	DATABASE CHEMICAL ABSTRACTS FILE SERVER STN KARLSRUHE ABSTRACT 119:21962, MINTON: "PHYSICAL CHARACTERIZATION OF CLOSTRIDIUM BOTULINUM NEUROTOXIN GENES" XP002015942 & REPORT (1992),ORDER NO.AD-A248904... see abstract ---	1-23
A	SYSTEM.APPL.MICROBIOL., vol. 18, no. 1, May 1995, pages 23-31, XP000605363 ELMORE ET AL: "NUCLEOTIDE SEQUENCE OF THE GENE CODING FOR PROTEOLYTIC (GROUP I) CLOSTRIDIUM BOTULINUM TYPE F NEUROTOXIN: GENEALOGICAL COMPARISON WITH OTHER CLOSTRIDIAL NEUROTOXINS" see the whole document ---	1-23
A	WO,A,94 21684 (PHLSB ;NIBSC (GB); SESARDIC DOROTHEA (GB); CHAN WOON LING (GB); SH) 29 September 1994 see page 12; table 1 see page 17,last paragraph-page 18,last paragraph see page 20; table 6 ---	1-23

-/--

2

C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	TRENDS IN BIOTECHNOLOGY, vol. 8, no. 4, April 1990, pages 88-93, XP000103110 SASSENFELD: "ENGINEERING PROTEINS FOR PURIFICATION" see the whole document ---	7,8,10, 12,13, 17-19, 21-23
P,X	BIOCHEMISTRY, vol. 34, November 1995, pages 15175-15181, XP002015939 ZHOU ET AL: "EXPRESSION AND PURIFICATION OF THE LIGHT CHAIN OF BOTULINUM NEUROTOXIN A: A SINGLE MUTATION ABOLISHES ITS CLEAVAGE OF SNAP-25 AND NEUROTOXICITY AFTER RECONSTITUTION WITH THE HEAVY CHAIN" see the whole document -----	7,8,10, 12,13, 17-19, 21-23

| Box I | Observations where certain claims were found unsearchable (Continuation of item 1 of first sheet) |

This International Search Report has not been established in respect of certain claims under Article 17(2)(a) for the following reasons:

1. [X] Claims Nos.: 22,23
because they relate to subject matter not required to be searched by this Authority, namely:

Remark: Although claims 22 and 23 are directed to a method of treatment of the human/animal body, the search has been carried out and based on the alleged effects of the compound/composition.

2. [ ] Claims Nos.:
because they relate to parts of the International Application that do not comply with the prescribed requirements to such an extent that no meaningful International Search can be carried out, specifically:

3. [ ] Claims Nos.:
because they are dependent claims and are not drafted in accordance with the second and third sentences of Rule 6.4(a).

| Box II | Observations where unity of invention is lacking (Continuation of item 2 of first sheet) |

This International Searching Authority found multiple inventions in this international application, as follows:

1. [ ] As all required additional search fees were timely paid by the applicant, this International Search Report covers all searchable claims.

2. [ ] As all searchable claims could be searches without effort justifying an additional fee, this Authority did not invite payment of any additional fee.

3. [ ] As only some of the required additional search fees were timely paid by the applicant, this International Search Report covers only those claims for which fees were paid, specifically claims Nos.:

4. [ ] No required additional search fees were timely paid by the applicant. Consequently, this International Search Report is restricted to the invention first mentioned in the claims; it is covered by claims Nos.:

Remark on Protest

[ ] The additional search fees were accompanied by the applicant's protest.

[ ] No protest accompanied the payment of additional search fees.

Patent document cited in search report	Publication date	Patent family member(s)		Publication date
WO-A-9403615	17-02-94	AU-A-	4719393	03-03-94
		CA-A-	2141427	17-02-94
		EP-A-	0652962	17-05-95
		FI-A-	950396	30-01-95
		JP-T-	8503602	23-04-96
		NO-A-	950348	28-03-95
WO-A-9421684	29-09-94	AU-A-	6432594	11-10-94
		CA-A-	2158748	29-09-94
		EP-A-	0690875	10-01-96

Possible Dissociation of the Heparin-binding and Mitogenic Activities of Heparin-binding (Acidic Fibroblast) Growth Factor-1 from Its Receptor-binding Activities by Site-directed Mutagenesis of a Single Lysine Residue

Wilson H. Burgess,* Anne M. Shaheen,* Mark Ravera,‡ Michael Jaye,‡ Patrick J. Donohue,* and Jeffrey A. Winkles*

* Laboratory of Molecular Biology, Jerome H. Holland Laboratory for the Biomedical Sciences, American Red Cross, Rockville, Maryland 20855; and ‡ Rorer Biotechnology, Inc., King of Prussia, Pennsylvania 19406

Abstract. The fibroblast or heparin-binding growth factors (HBGFs) are thought to be modulators of cell growth and migration, angiogenesis, wound repair, neurite extension, and mesoderm induction. A better understanding of the structural basis for the different activities of these proteins should facilitate the development of agonists and antagonists of specific HBGF activities and identification of the signal transduction pathways involved in the mechanisms of action of these growth factors. Chemical modification studies of Harper and Lobb (Harper, J. W., and R. R. Lobb. 1988. *Biochemistry.* 27:671–678) implicated lysine 132 in HBGF-1 (acidic fibroblast growth factor) as being important to the heparin-binding, receptor-binding, and mitogenic activities of the protein. We changed lysine 132 to a glutamic acid residue by site-directed mutagenesis of the human cDNA and expressed the mutant protein in *Escherichia coli* to obtain sufficient quantities for functional studies. Replacement of this lysine with glutamic acid reduces the apparent affinity of HBGF-1 for immobilized heparin (elutes at 0.45 M NaCl vs. 1.1 M NaCl for wild-type). Mitogenic assays established two points: (a) human recombinant HBGF-1 is highly dependent on the presence of heparin for optimal mitogenic activity, and (b) the change of lysine 132 to glutamic acid drastically reduces the specific mitogenic activity of HBGF-1. The poor mitogenic activity of the mutant protein does not appear to be due to a reduced affinity for the HBGF receptor. Similarly, the mutant HBGF-1 can stimulate tyrosine kinase activity and induce protooncogene expression. Differences in the biological properties of the wild-type and mutant proteins were observed in transfection studies. Mutant HBGF-1 expression in transfected NIH 3T3 cells did not induce the same transformed phenotype characteristic of cells expressing wild-type HBGF-1. Together these data indicate that different functional properties of HBGF-1 may be dissociated at the structural level.

1. *Abbreviation used in this paper: HBGF, heparin-binding growth factor.*

34). Intact growth factor per... intracellularly for several hours and large fragments (15,000 and 10,000 M_r for HBGF-1; 16,000 M_r for HBGF-2) are detectable after as many as 24 h. Further, nuclear or nucleolar localization of HBGF-2 has been observed (2, 35).

Despite the identification of additional members of the HBGF family and a broad range of cells and tissues that contain the growth factors, and despite the availability of large quantities of recombinant protein and increased knowledge of the broad spectrum of activities of potential biological significance that can be attributed to the HBGFs, relatively little is known regarding the relationship of these highly conserved structures to any of their known functions. Baird et al. (1) reported the synthesis of 25 peptides, which together encompass and overlap the entire sequence of HBGF-2 as described by Ueno et al. (42). They reported the identification of two functional domains in the primary structure of HBGF-2 based on the abilities of synthetic peptides to interact with HBGF receptor, bind radiolabeled heparin in a solid phase assay, and inhibit HBGF-2 stimulation of thymidine incorporation into DNA. Using the numbering system of the authors (which does not correspond to full length HBGF-2) statistically significant functional activities could be assigned to peptides corresponding to residues 24–68 and 106–115 of HBGF-2. Similarly, Schubert et al. (39) demonstrated that peptides corresponding to residues 1–24, 24–68, and 93–120 of HBGF-2 are able to stimulate substratum adhesion of PC12 cells. We have shown that a synthetic peptide corresponding to residues 49–72 of HBGF-1 (using numbering of 1–154 for full length HBGF-1) is able to compete with HBGF-1 for heparin binding in a gel overlay assay (33). This region is homologous to one of the regions of HBGF-2 (residues 24–68) described above as possessing heparin-binding activity.

To date, the most complete and informative studies documenting the effects of chemical modification of any HBGF on function are those of Harper and Lobb (19). Briefly, they were able to show that limited reductive methylation of bovine HBGF-1 with formaldehyde and cyanoborohydride resulted in stoichiometric methylation only of lysine 132 (using 1–154 numbering for full length HBGF-1). They reported 90% modification of this residue, with 60% dimethyllysine. The modified protein exhibited significantly reduced apparent affinity for immobilized heparin (eluted at ~0.7 M NaCl vs. ~1.2 M NaCl for unmodified HBGF-1), a fourfold reduction in its ability to stimulate DNA synthesis in NIH 3T3 fibroblasts and a similar reduction in its ability to compete with labeled ligand in a radioreceptor assay. A lysine residue is found at this position of HBGF-1 and HBGF-2 of all species characterized to date. Together these data implicate a crucial role for lysine 132 in several of the known functions of HBGF-1.

In this report we address the role of lysine 132 in HBGF-1 function using site-directed mutagenesis of this position to a glutamic acid. This approach offers several advantages over chemical modification studies including (a) the ability to produce large quantities of the desired product, (b) elimination of significant (although sub-stoichiometric) modification of other lysines, and (c) allowing the introduction of modified HBGF-1 into mammalian cells through transfection of cDNA expression vectors designed to produce the desired mutant. Despite these advantages the importance of chemi-

cal modification studies such as those of Harper and Lobb (19) should not be underestimated for they are extremely useful in the design of a rational approach to site-directed mutagenesis. The results described here demonstrate that replacement of lysine 132 ... ...with glutamic acid reduces significantly its apparent affinity for immobilized heparin and its mitogenic capacity. However, the apparent affinity of the mutant for high affinity cell surface receptors appears unaltered. When assayed in the presence of heparin where the difference in wild-type and mutant HBGF-1 mitogenic activity is most apparent, mutant HBGF-1 can stimulate tyrosine kinase activity and induce protooncogene expression. Functional differences between the wild-type and mutant HBGF-1 are also apparent after transfection of cDNA expression vectors into NIH 3T3 fibroblasts.

Materials and Methods

Materials

Heparin-Sepharose, protein A-Sepharose, pKK233 expression vectors, and low molecular weight markers were purchased from Pharmacia Fine Chemicals (Piscataway, NJ). All reagents for PAGE and the Mighty Small Apparatus were from Hoefer Scientific Instruments (San Francisco, CA). Reagents for reversed-phase HPLC, amino acid analysis, and amino acid sequencing were purchased from Applied Biosystems, Inc. (Foster City, CA). Isotopes and the in vitro mutagenesis system were from Amersham Corp. (Arlington Heights, IL). The rabbit polyclonal HBGF-1-specific antibody was provided by R. Friesel (American Red Cross, Rockville, MD) and the rabbit polyclonal anti-phospholipase C-γ antibodies were provided by A. Zilberstein (Rorer Biotechnology, Inc., King of Prussia, PA). Tissue culture media and plasticware were purchased from Gibco Laboratories (Grand Island, NY). High molecular weight molecular markers were from Bio-Rad Laboratories (Richmond, CA). Endoproteinase ASP-N and the random primer DNA labeling kit were from Boehringer Mannheim Biochemicals (Indianapolis, IN). Other chemicals were reagent grade.

Construction of pREC and pl32E Prokaryotic Expression Plasmids

The plasmid expressing wild-type HBGF-1 (corresponding to the α-form of endothelial cell growth factor (5), pREC, was kindly provided by R. Forough (American Red Cross). This plasmid was constructed by cloning synthetic oligonucleotide cassettes into the Nco I/Hind III site of pKK233-2. The plasmid expressing mutant HBGF-1 (glutamic acid instead of lysine at amino acid position 132; pl32E) was constructed as follows. The Eco RI/Hind III fragment of HBGF-1 cDNA clone 1 (2.0) was subcloned into M13mp18. Single-stranded template was prepared and used for oligonucleotide-directed in vitro mutagenesis. Double-stranded DNA was transformed into E. coli TG-1 cells and the resultant plaques were screened by M13 dideoxy sequencing. The mutated HBGF-1 cDNA was transferred into the expression vector pKK223-3 using the original Eco RI and Hind III sites.

Production and Purification of Recombinant Proteins

Recombinant plasmids pREC or pl32E were introduced into the lacI^q-bearing Escherichia coli strain JM103. Cultures of JM103 bearing the recombinant plasmids were grown with shaking at 37°C in Luria broth containing 100 μg/ml ampicillin. A fresh overnight culture was diluted and grown until the A$_{550}$ reached ~0.2, at which point isopropylthio-β-galactoside was added to 1 mM. Cells were collected by centrifugation and from ~ -80°C for subsequent growth factor purification.

The ... en cell pellets from 2-liter cultures were resuspended in 50 ml of 10 mM Tris-HCl, pH 7.5, 5 mM EDTA, 50 mM glucose. A fresh solution of hen egg lysozyme in the same buffer was added to 10 μg/ml. The cells were mixed at 4°C for 45 min. The viscous lysate was sonicated at maximum intensity using a large probe and four 20-s pulses of a Heat Systems W-380 sonicator. The lysate was clarified by centrifugation at 6,000 g for 15 min at 4°C. The supernatant was diluted to 100 ml with 50 mM Tris-HCl, pH 7.5, 10 mM EDTA and incubated with 20 ml of hydrated heparin-

Sepharose at 4°C with end-over-end mixing for 2 h. The resin was eluted batchwise using a sintered glass funnel and successive washes of the same buffer containing 0, 0.1, 0.5, 0.65, and 1.5 M NaCl.

The wild-type recombinant HBGF-1 eluted with the 1.5 M NaCl wash. The mutant was eluted with the 0.5 M NaCl wash. Although the wild-type protein was essentially pure after heparin-Sepharose chromatography, the mutant HBGF-1 constituted only 10–20% of the 0.5 M NaCl wash. Both preparations were purified to >95% purity using reversed-phase HPLC (4). The reversed-phase purified material was used for all reported studies.

Characterization of Recombinant Proteins

All preparations of purified recombinant human wild-type and mutant HBGF-1 were analyzed by SDS-PAGE, amino acid analysis, amino terminal sequencing, peptide mapping, and amino acid sequencing of the peptide encompassing the mutated residue. Protein concentrations were determined by amino acid analysis. Aliquots of wild-type and mutant HBGF-1 were subjected to electrophoresis using the SDS-PAGE system of Laemmli (26). A 15% acrylamide, 0.4% N,N-methylenebisacrylamide solution was polymerized in a Hoefer mini-gel apparatus and electrophoresis was carried out at a constant 200 V. Protein was visualized by staining the gel with 0.1% Coomassie blue R-250 in 50% methanol, 10% glacial acetic acid, and destaining with 9% glacial acetic acid, 5% methanol. Samples for amino acid analysis were hydrolyzed with argon-purged, constant boiling 6 N HCl at 115°C for 18 h using a Pico-Tag workstation (Waters Associates, Milford, MA). Amino acids were derivatized with phenylisothiocyanate and separated with a PTC analyzer (model 130A; Applied Biosystems, Inc.). A Waters 840 system was used for data collection and reduction. Amino acid sequences were established using a protein sequencer (model 477A; Applied Biosystems, Inc.) using modified Edman chemistry and an on-line model 120A PTH analyzer. Peptide mapping of recombinant protein after digestion with endoproteinase Asp-N at a 1:25 ratio of enzyme to protein in 50 mM Na₂HPO₄, pH 8.0, 37°C for 18 h was performed using a microbore HPLC system (model 130A; Applied Biosystems, Inc.). The appropriate peptides were subjected to amino acid sequence analysis to establish the fidelity of expression of the wild-type and mutant HBGF-1 vectors.

Stability Studies

Metabolically labeled recombinant proteins were prepared by growing bacterial cultures as described above until the A₅₅₀ reached ~0.4, at which point the cells were collected by centrifugation. They were resuspended in 98.5% M9 minimal medium/1.5% Luria broth and [³H]leucine (140 Ci/mmol) was added to 45 μCi/ml. Cells were grown with shaking for 30 min, and then for an additional 4 h in the presence of 1 mM isopropylthio-β-galactoside. Cells were collected and growth factors purified as described above. The purified labeled growth factors were incubated for 48 h at 37°C in the presence of media (DMEM containing 10% calf serum) that had been conditioned for 48 h by NIH 3T3 cells. The growth factor-containing media was analyzed by SDS-PAGE and autoradiography.

Mitogenic Assays

The mitogenic activities of wild-type and mutant recombinant HBGF-1 were determined by measuring their ability to stimulate DNA synthesis in NIH 3T3 cells and to support the proliferation of human umbilical vein endothelial cells. DNA synthesis was determined by measuring the amount of [³H]thymidine incorporated into cells. Briefly, NIH 3T3 cells were seeded into 48-well plates and grown to near confluence in DME containing 10% calf serum. The cells were serum starved (DME, 0.5% serum) for 24 h. Mitogens were added to the wells and incubated for 18 h. The cells were pulsed with 0.5 μCi/ml of [³H]thymidine (25 Ci/mmol) for 4 h. The cells were rinsed with PBS, fixed with 10% TCA, rinsed with PBS, and then solubilized with 0.5 N NaOH. Incorporation of [³H]thymidine into acid-insoluble material was determined by scintillation counting. All assays were performed in triplicate.

Human umbilical vein endothelial cells were provided by T. Maciag (American Red Cross, Rockville, MD). They were maintained on fibronectin-coated plates (2 μg/cm²) in medium 199 supplemented with 10% (vol/vol) heat-inactivated FBS, 1× antibiotic-antimycotic, 10 U/ml heparin, and 10 ng/ml human recombinant HBGF-1. For growth assays, cells were seeded in 24-well plates at 2,000 cells/well in medium 199 supplemented as above with the exception of HBGF-1. The indicated amounts of wild-type or mutant HBGF-1 and heparin were added to the wells. The media was changed every other day. After 7 d in culture, cells were trypsinized and counted using a hemocytometer.

Competition for Binding and Cross-Linking to Cell Surface Receptors

Bovine brain-derived HBGF-1 (4) was labeled with ¹²⁵I using immobilized lactoperoxidase and biologically active, labeled protein was isolated using heparin-Sepharose as described (16). Confluent NIH 3T3 cells in 24-well plates were serum starved for 24 h before binding experiments in DME containing 0.5% calf serum. The cells were washed and incubated with DME containing 5 U/ml heparin, 0.5% BSA, and 25 mM Hepes, pH 7.2 (binding buffer) at room temperature for 20 min. The cells were then incubated with ¹²⁵I-HBGF-1 and unlabeled wild-type or mutant HBGF-1 in the presence of 5 U/ml heparin as indicated in the figure legend. The cells were incubated on ice for 90 min. The plates were aspirated and washed four times with binding buffer. The cells were then incubated for 20 min at 4°C with 1 ml of 0.3 mM disuccinimidyl suberate in PBS. The cross-linker was then aspirated off and the reaction quenched by adding 2.0 M Tris-HCl, pH 8.0. The cells were washed with PBS, scraped from the plates and pelleted for 10 s at 15,000 g. The pellet were extracted with 100 μl of 50 mM Tris, 1 mM EDTA, 100 mM NaCl, 1.0% Triton X-100, 0.1 mM phenylmethylsulfonyl fluoride, pH 7.5 for 20 min at 4°C. The extracts were centrifuged for 10 min at 15,000 g. The supernatants were removed and mixed with an equal volume of Laemmli sample buffer for SDS-PAGE analysis.

Stimulation of Protein Tyrosine Kinase Activity

NIH 3T3 cells were grown to confluence in 100 mm dishes and serum starved as described above. The cells were then exposed to diluent 1.0, or 10 ng/ml of wild-type or mutant HBGF-1 for 10 min at 37°C. The cells were washed once with cold PBS then lysed in buffer containing 10 mM Tris, 50 mM NaCl, 5 mM EDTA, 50 mM NaF, 30 mM sodium pyrophosphate, 100 μM sodium orthovanadate, 1.0% Triton X-100, 1 mM phenylmethylsulfonyl fluoride, pH 7.4. The cells were scraped from the plates, vortexed, and incubated on ice for 10 min. Lysates were clarified by centrifugation at 10,000 g for 10 min at 4°C and the supernatants were mixed with an equal volume of 2× Laemmli sample buffer. Samples (normalized to cell number) were subjected to PAGE in the presence of SDS. The proteins were transferred to nitrocellulose and immunoblotted with anti-phosphotyrosine antibodies as described (15). The blots were incubated with ¹²⁵I-protein A and phosphotyrosine-containing proteins were visualized by autoradiography. In some experiments the initial cell lysates were incubated with a pre-bound anti-phospholipase C-γ antibody/protein A-Sepharose complex (30) for 90 min at 4°C. The beads were washed with 20 mM Hepes, 0.1% Triton X-100, 150 mM NaCl, 10% glycerol, pH 7.5. Immunoprecipitated proteins were eluted from the beads with 2× Laemmli sample buffer and subjected to PAGE and Western blotting with anti-phosphotyrosine antibodies as described above.

RNA Gel Blot Analysis

NIH 3T3 cells were incubated for 48 h in DME/0.5% FCS and then either left unstimulated or stimulated with wild-type or mutant HBGF-1 for the indicated times. Cells were harvested, total RNA was prepared (17), and 10 μg of each sample was separated by electrophoresis on 1.2% agarose gels containing formaldehyde. The gels were stained with ethidium bromide and photographed to verify that each lane contained an equal amount of undegraded ribosomal RNA. RNA was electroblotted onto Zetabind nylon filters and cross-linked by UV irradiation. The restriction fragments used and source of the DNA probes were as follows: (a) c-fos, 2.8-kb Nco I/Xho I fragment of pc-fos-1; American Type Culture Collection, Rockville, MD; (b) c-jun, 1.5-kb Hind III/Bam HI fragment of ph-cJ-1; gift of P. Angel, University of California, La Jolla, CA; (c) c-myc, 1.4-kb Sst I fragment of pHSR-1; ATCC; (d) glyceraldehyde 3-phosphate dehydrogenase, 0.8-kb Pst I/Xba I fragment of pHcGAP; ATCC. The probes were labeled with [³²P]dCTP (3,000 Ci/mmol) using a random primer labeling kit. Hybridization and filter washes were as described (17). Blots were exposed to Kodak XAR5 film at −70°C.

Transfection of NIH 3T3 Cells with HBGF-1 Eukaryotic Expression Plasmids

NIH 3T3 cells in 100 mm dishes were transfected with plasmid DNA by the calcium phosphate precipitation method (44). Cells were incubated with either 1 μg of pSV2 neo (41) or co-transfected with a mixture (1:10 μg) of pSV2 neo and either HBGF-1 wild-type expression vector (p267) or HBGF-1 mutant expression vector (p268). The plasmid p267 is described in Jaye et

Burgess et al. Site-directed Mutagenesis of HBGF-1

2131
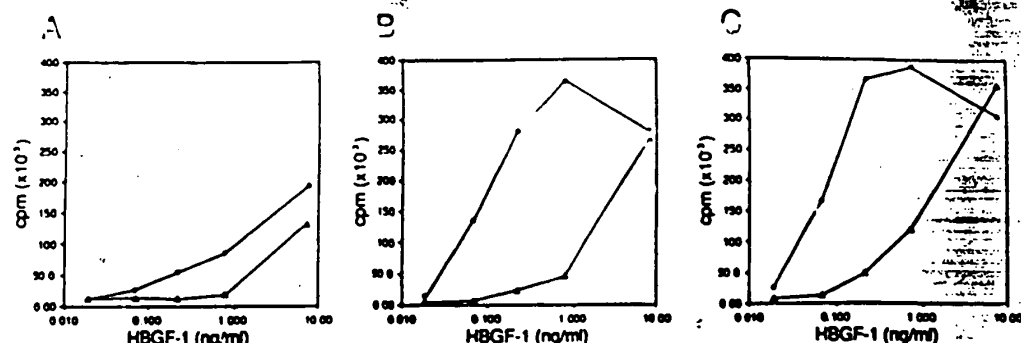
Figure I. Stimulation of DNA synthesis in NIH 3T3 cells by wild-type and mutant HBGF-1. Cells were grown to near confluence and serum starved for 24 h as described in Materials and Methods. Cells were treated with the indicated concentrations of wild-type (●) or mutant (▲) HBGF-1, incubated for 18 h, and then pulsed with 0.5 μCi of [³H]thymidine/ml for 4 h. The cells were harvested and incorporation of radioactivity was determined. Both wild-type and mutant HBGF-1 were assayed in the presence of 0 (*A*), 5 (*B*), or 50 U/ml heparin (*C*).

al. (23); p268 was constructed by replacing the 297nt Pvu II/Bgl II fragment of p267 (encoding amino acids 38–155) with the corresponding region from the prokaryotic expression plasmid pE132 using standard subcloning methods. Cells were split to 10 dishes and transfected colonies were selected by incubating the cells in DME, 10% calf serum containing 500 μg/ml Geneticin. The media was changed every 3–4 d. After 4 wk, transfected colonies were analyzed for HBGF-1 expression by Western blot analysis using rabbit polyclonal HBGF-1-specific antibodies and ¹²⁵I-protein A as described above.

Results

Heparin-binding Properties of HBGF-1 Mutant pl32E

A drastic reduction in the apparent affinity of HBGF-1 containing glutamic acid in place of lysine at position 132 was observed during the purification of the recombinant proteins

Figure 2. Ability of wild-type and mutant HBGF-1 to stimulate growth of human umbilical vein endothelial cells. Cells were seeded and cultured as described in Materials and Methods. Cell number after 7 d in culture in the presence of the indicated concentrations of wild-type (○/●) or mutant (△/▲) HBGF-1 in the absence (○/△) or presence (●/▲) of 50 U/ml heparin is shown.

from the *Escherichia coli* lysates. Recombinant wild-type HBGF-1 from *E. coli* lysates can be purified t near homogeneity with a single heparin-Sepharose step. The protein binds the immobilized heparin during extensive washing with 0.5 and 0.65 M NaCl-containing buffers and is eluted with a single step of 1.5 M NaCl-containing buffer. In contrast, heparin-Sepharose affinity-based chromatography could not be used as a single purification step for the mutant HBGF-1. The mutant protein binds immobilized heparin in the presence of 0.1 M NaCl but was eluted during the 0.5 M NaCl wash. Both wild-type and mutant HBGF-1 (1.5 and 0.5 M NaCl eluates, respectively) could be purified to apparent homogeneity using reversed-phase HPLC. Detailed analysis of the apparent affinities of the two purified proteins for immobilized heparin-Sepharose using relatively shallow, linear NaCl gradients indicated that the mutant HBGF-1 eluted with 0.45 M NaCl whereas wild-type required 1.1 M NaCl to be eluted (data not shown). For all of the assays described below we used reversed-phase HPLC purified wild-type or mutant HBGF-1. Protein concentrations were determined by amino acid analysis of preparations that had been shown to be the desired HBGF-1 form by peptide mapping and amino acid sequence analysis (data not shown).

Mitogenic Properties of HBGF-1 Mutant pl32E

The ability of the HBGF-1 mutant to stimulate mitogenesis was compared to that of the wild-type protein using two different assays. In the first, the ability of the two proteins to stimulate DNA synthesis in NIH 3T3 cells as measured by [³H]thymidine incorporation was examined. The assays were conducted over a broad range of HBGF-1 and heparin concentrations. Two important points can be made from the data in Fig. 1. One, the wild-type HBGF-1 has a dramatic requirement for the presence of heparin for optimal mitogenic activity and, two, the mutant HBGF-1 is significantly less potent than wild-type protein in the presence of added heparin. As can be seen in Fig. 1, the maximal difference in mitogenic potency was observed in the presence of 5 U/ml heparin (~30-fold). Little difference (approximately three-

Table I. Cell Number (× 10⁻³)

	Growth factor concentration (ng/ml)					
	0	0.1	0.5	1	5	10
GLU₁₃₂ HBGF-1	1.6	1.6	1.3	1.2	1.7	1.4
Wild-type HBGF-1	1.7	2.0	1.9	2.9	12.6	16.6

fold) between the wild-type and mutant protein is seen in the absence of added heparin because of the relative lack of mitogenic activity of wild-type human recombinant HBGF-1 in the absence of heparin. The possibility that the reduced mitogenic activity of the mutant HBGF-1 is related directly to its reduced apparent affinity for immobilized heparin is supported by the observation that the difference in the mitogenic potency between the wild-type and mutant protein is reduced to ~18-fold in the presence of 50 U/ml heparin.

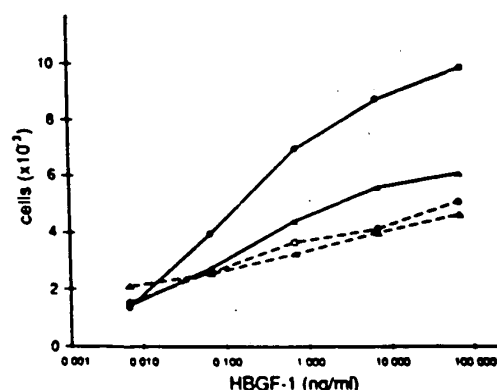In the second mitogenesis assay the abilities of the wild-type and mutant proteins to support the proliferation of human umbilical vein endothelial cells were compared. The results shown in Fig. 2 are consistent with those described above in that they demonstrate a dramatic heparin requirement of the wild-type HBGF-1 for biological activity and that the mutant HBGF-1 is not able to support cell proliferation to the same extent as the wild-type protein. These experiments were conducted in the presence of 50 U/ml heparin and the endothelial cells were seeded in the presence of 10 ng/ml wild-type HBGF-1. When growth assays were conducted in the presence of 5 U/ml heparin without wild-type protein during the seeding, mitogenic deficiencies of the mutant protein were more pronounced (Table I). The results shown in Fig. 3 demonstrate that the reduced mitogenic activity of the mutant HBGF-1 does not appear to be the result of any increased susceptibility of the protein to proteolytic digestion by components in serum or the conditioned media of NIH 3T3 cells.

Receptor-binding Activity of HBGF-1 Mutant pL32E

The results presented above are consistent with the observa-

Figure 1. Analysis of the relative stability of wild-type and mutant HBGF-1 in NIH 3T3 cell-conditioned media. The wild-type and mutant proteins were labeled and purified as described in Materials and Methods. The proteins were incubated in the presence of NIH 3T3 cell-conditioned media for 48 h at 37°C and then subjected to SDS-PAGE. The gels were dried and labeled proteins visualized by autoradiography. Lane 1 contains wild-type HBGF-1 and lane 2 mutant HBGF-1. The apparent molecular weights of both proteins are identical to that of HBGF-1 before incubation.

tions of Harper and Lobb (19) using bovine brain-derived HBGF-1 selectively methylated at lysine 132, although the magnitude of the reduction in mitogenic potency (~30-fold for 3T3 cell assay) as compared with the ~4-fold decrease reported by *‡——* and Lobb (19) is significantly greater. They also reported r⋯⋯ ⋯⋯⋯-binding activity for the modified protein. We examined the abilities of the wild-type and mutant recombinant HBGF-1 to compete with ¹²⁵I-labeled bovine HBGF-1 for binding to cell surface receptors on NIH 3T3 cells at a concentration of added heparin (5 U/ml) where the difference in mitogenic potencies of the two proteins was greatest.

The receptor-binding activity of the mutant HBGF-1 was established by competition for cross-linking of ¹²⁵I-HBGF-1 to 150,000- and 130,000-M, proteins present on the surface of NIH 3T3 cells (16). The results shown in Fig. 4 demonstrate that the mutant HBGF-1 is similar to wild-type protein in its ability to compete for receptor-ligand cross-linking.

The functional consequences of HBGF-1 binding to its cell surface receptor include stimulation of protein tyrosine kinase activity (8, 15, 20) including phosphorylation of phospholipase C-γ (6). Fig. 5 A demonstrates that both wild-type and mutant HBGF-1 are able to increase the phosphotyrosine content of 150,000-, 90,000-, and 70,000-M, proteins and, to a lesser extent, proteins with lower relative molecular masses as judged by Western blot analysis with phosphotyrosine-specific antibodies. The dose response and extent of activation is similar for the two forms of the growth factor. Stimulation of the phosphotyrosine content of phospholipase C-γ was examined by anti-phosphotyrosine Western blot analysis of 3T3 cell lysates after immunoprecipitation using antibodies that recognize phospholipase C-γ. Fig. 5 B demonstrates that mutant HBGF-1 shares with wild-type HBGF-1 the ability to stimulate tyrosine phosphorylation of phospholipase C-γ. These data regarding stimulation of tyrosine kinase activity by wild-type and mutant HBGF-1 are in good agreement with the receptor-binding data described above but do not provide insight into the functional basis for the relatively poor mitogenic capacity of this HBGF-1 mutant.

Protooncogene Induction by Wild-Type and Mutant HBGF-1

The results described above indicate that the functional properties of the mutant HBGF-1 associated with events that occur at the cell surface (i.e., receptor-binding and tyrosine kinase activation) are normal with respect to those of wild-type HBGF-1. In addition to tyrosine kinase activation, another early response to HBGF-1 receptor-binding is the elevation of protooncogene mRNA levels (17). To determine the effect of wild-type and mutant HBGF-1 on protooncogene expression, NIH 3T3 cells were serum starved and then either left unstimulated or stimulated with 10 ng/ml wild-type or mutant HBGF-1. Heparin (5 U/ml) was also added to the cells receiving growth factor. Cells were collected at various times after stimulation, RNA was prepared, and levels of c-fos, c-jun, c-myc, and glyceraldehyde 3-phosphate dehydrogenase mRNA (as a control for the amount of RNA loaded in each lane) were assayed by RNA gel blot analysis. Wild-type and mutant HBGF-1 increased protooncogene mRNA levels to a similar degree; maximal levels were observed at 30 min (c-fos, c-jun) or 2 h (c-myc) after stimulation (Fig.

150 ►
130 ►

1 2 3 4 5 6 7 8 9 10

Figure 4. Ability of wild-type and mutant HBGF-1 to compete with ^{125}I-labeled bovine HBGF-1 for cross-linking to 150,000- and 130,000-mol wt cell surface receptors. NIH 3T3 cells were incubated with 1 ng/ml bovine ^{125}I-HBGF-1 and either 0.5, 1.0, 5.0, 10.0, or 50.0 ng/ml of wild-type (lanes 1-5) or mutant (lanes 6-10) human recombinant HBGF-1 in the presence of 5 U/ml heparin. After incubation, the cells were treated with cross-linking reagents as described in Materials and Methods. The apparent molecular weights of cross-linked species were determined after SDS-PAGE and autoradiography. The positions of two cross-linked 150,000- and 130,000-mol wt species, which correspond to the known apparent molecular weights of HBGF receptors, are indicated with arrows.

6). The addition of heparin alone did not induce protooncogene expression. Since the mitogenic differences between the wild-type and mutant HBGF-1 are more pronounced at lower growth factor concentrations, we also stimulated cells with 0.5, 1.0, 5.0, and 10.0 n' of wild-type and mutant growth factor (again in the presence of heparin). At all four concentrations used, the wild-type and mutant HBGF-1 were similar in their ability to induce c-*fos* mRNA expression (Fig. 7).

Overexpression of Wild-Type and Mutant HBGF-1 in Transfected NIH 3T3 Cells

It was demonstrated previously that overexpression of wild-type HBGF-1 in transfected Swiss 3T3 cells resulted in cells with an elongated, transformed morphological phenotype that grew to high saturation densities (23). This transformed phenotype occurred even though the HBGF-1 polypeptide was not detectable in the conditioned media of these cells. We have shown that the mutant HBGF-1 is not a potent mitogen although it can bind receptor and initiate early events associated with mitogenic signal transduction. To investigate whether the intracellular function of the mutant HBGF-1 was altered, we examined the ability of this protein to induce a transformed phenotype in NIH 3T3 cells. Cells were either transfected with a plasmid conferring neomycin resistance or co-transfected with the neomycin resistance plasmid and wild-type or mutant HBGF-1 expression vectors.

Fig. 8 shows the results of Western blot analysis of transfected cell lysates using HBGF-1-specific antibodies. The Western blot analysis was normalized to cell number and provides the basis for our designation of relatively high or low levels of HBGF-1 expression. The results shown in Fig. 9 demonstrate that cells expressing a high level of wild-type HBGF-1 (Fig. 9 *B*) and to some extent a low level of wild-

A   **B**

200 ►
115 ►
94 ►
67 ►
43 ►
30 ►

1 2 3 4 5   1 2 3

◄ PLC

Figure 5. Stimulation of protein tyrosine kinase activity by wild-type and mutant HBGF-1. (*A*) Serum starved NIH 3T3 cells were either (lane *1*) unstimulated or treated with 5 U/ml heparin and (lane *2*) 1 ng/ml wild-type; (lane *3*) 10 ng/ml wild-type; (lane *4*) 1 ng/ml mutant; or (lane *5*) 10 ng/ml mutant HBGF-1. The cells were processed as described in Materials and Methods and phosphotyrosine-containing proteins were visualized using antiphosphotyrosine antibodies and ^{125}I-protein A. The arrows indicate the positions of 150,000-, 90,000-, and 70,000-mol wt proteins whose phosphotyrosine content are increased by the addition of wild-type or mutant HBGF-1. (*B*) Cells were incubated as in *A* with the exception that cell lysates were immunoprecipitated with anti-phospholipase C-γ antibodies before Western blot analysis with anti-phosphotyrosine antibodies. Cells were either (lane *1*) unstimulated or treated with (lane *2*) 10 ng/ml wild-type, or (lane *3*) 10 ng/ml mutant HBGF-1. The arrow shows the position of a 150,000-mol wt protein whose phosphotyrosine content is increased by treatment with wild-type or mutant HBGF-1.

Figure 6. Effect of wild-type and mutant HBGF-1 on protooncogene mRNA levels. Serum starved NIH 3T3 cells were either left un-stimulated or treated with heparin (5 U/ml) and 10 ng/ml wild-type (wt) or mutant (mt) HBGF-1 for the indicated time periods. RNA was prepared and used for RNA gel blot hybridization using the radiolabeled DNA probes indicated on the left side (GAPDH, glyceraldehyde 3-phosphate dehydrogenase). The upper and lower tick marks on the left side of each panel represent the positions of 28 and 18S rRNA, respectively.

stoichiometric modification of lysine residue 132 (using the 1-154 numbering system for full-length HBGF-1). It was suggested that modification of this residue, which is conserved in all HBGF-1 and HBGF-2 sequences reported to date, was responsible for the reduced affinity. ^nity for immobilized heparin, the reduced mitogenic capacity, and the reduced receptor-binding activity of the modified protein. The results presented here using site-directed mutagenesis to address the role of lysine 132 on the functional properties of HBGF-1 are in general agreement with the conclusions of Harper and Lobb (19). Specifically, substitution of lysine 132 for glutamic acid reduces the apparent affinity of the recombinant protein for immobilized heparin (elutes at 0.45 M NaCl compared with 1.1 M NaCl for wild-type) and significantly reduces the mitogenic potency of the growth factor. The reduced mitogenic potency may be a direct consequence of the reduced apparent affinity of the mutant HBGF-1 for heparin since it has been demonstrated that the class 1 heparin-binding growth factors in general (29) and human HBGF-1 in particular (22, 43) are dependent on the presence of heparin for optimal biological activity.

Our results do not support the notion that the reduced mitogenic capacity of HBGF-1 containing glutamic acid in place of lysine at position 132 is due to reduced binding to cell surface receptors. The receptor-binding properties of the mutant HBGF-1 are not distinguishable from those of the wild-type protein as judged by cross-linking experiments (see Fig. 4). In addition, the mutant HBGF-1 is able to induce the same pattern of tyrosine kinase phosphorylation as is the wild-type protein (see Fig. 5) and can induce proto-oncogene expression (see Fig. 6). The majority of the studies presented here utilize a heparin concentration of 5 U/ml; the concentration where maximal difference between the mitogenic activity of wild-type and mutant HBGF-1 was observed in the 3T3 cell thymidine incorporation assay. It should be noted that in the absence of heparin, the mutant HBGF-1 competes poorly with labeled wild-type HBGF-1 in cross-

type HBGF-1 (Fig. 9 D) have acquired a more polar, elongated phenotype characteristic of transformed 3T3 cells. This phenotype is not seen in cells expressing neomycin resistance alone (Fig. 9 A) or in cells expressing relatively high levels of mutant HBGF-1 (Fig. 9 C). It should be noted that we have not been able to detect HBGF-1 immunoreactivity in the media conditioned by these cells and that the cells expressing relatively high levels of wild-type HBGF-1 show enhanced growth in soft agar relative to untransfected cells or cells expressing high levels of the mutant HBGF-1 (data not shown). These results are consistent with the results of the mitogenic assays described above which demonstrate that the growth-promoting activity of the mutant HBGF-1 is relatively low when compared to the wild-type protein.

Discussion

The experiments described in this report were initiated as a result of the chemical modification studies of HBGF-1 reported by Harper and Lobb (19). They demonstrated that reductive methylation of HBGF-1 resulted in selective,

Figure 7. Effect of different concentrations of wild-type and mutant HBGF-1 on c-fos mRNA levels. Serum starved NIH 3T3 cells were either left un-stimulated or treated with heparin (5 U/ml) and (A) 0.5 ng/ml, (B) 1.0 ng/ml, (C) 5.0 ng/ml, (D) 10 ng/ml wild-type (wt) or mutant (mt) HBGF-1 for the indicated time periods. RNA was prepared and used for RNA gel blot hybridization using the c-fos DNA probe (upper panels) or glyceraldehyde 3-phosphate dehydrogenase DNA probe (lower panels).

1 2 3 4 5

linking assays (data not shown). In addition, whereas the apparent affinity of the mutant HBGF-1 for immobilized heparin is reduced, it does bind at ionic strengths (i.e., ~0.5 M NaCl) that exceed those known to be physiologic. Thus, the data presented here indicate that the mutant can utilize the

presence of heparin to restore some (i.e., receptor-binding, tyrosine kinase activation, and protooncogene induction) but not all (i.e., stimulation of [^{3}H]thymidine incorporation into DNA and endothelial cell proliferation) of the activities of the wild-type pro... ...rly, it is of interest that the wild-type protein competes with label d HBGF-1 for receptor-binding and induces protooncogene expression at similar concentrations in the presence or absence of added heparin yet it requires added heparin in order to promote DNA synthesis and cell proliferation (Figs. 1, 2, 4, and 6; and data not shown). Thus, the relatively poor mitogenic activity of the mutant protein may be related to its reduced apparent affinity for heparin. The data presented here demonstrate that "high" affinity receptor-binding, activation of tyrosine kinase activity, tyrosine phosphorylation of specific substrates, and in:h:'"ion of protooncogene expression may be necessary but are not, by themselves, sufficient to sustain a mitogenic response to the presence of HBGF-1. These results are consistent with the observations of Escobedo and Williams (12) who showed by site-directed mutagenesis of the PDGF receptor and cDNA transfection that mutants could be construc:.d that were responsive to PDGF with respect to receptor tyrosine kinase activation and increased phosphatidylinositol turnover but did not elicit a mitogenic re-
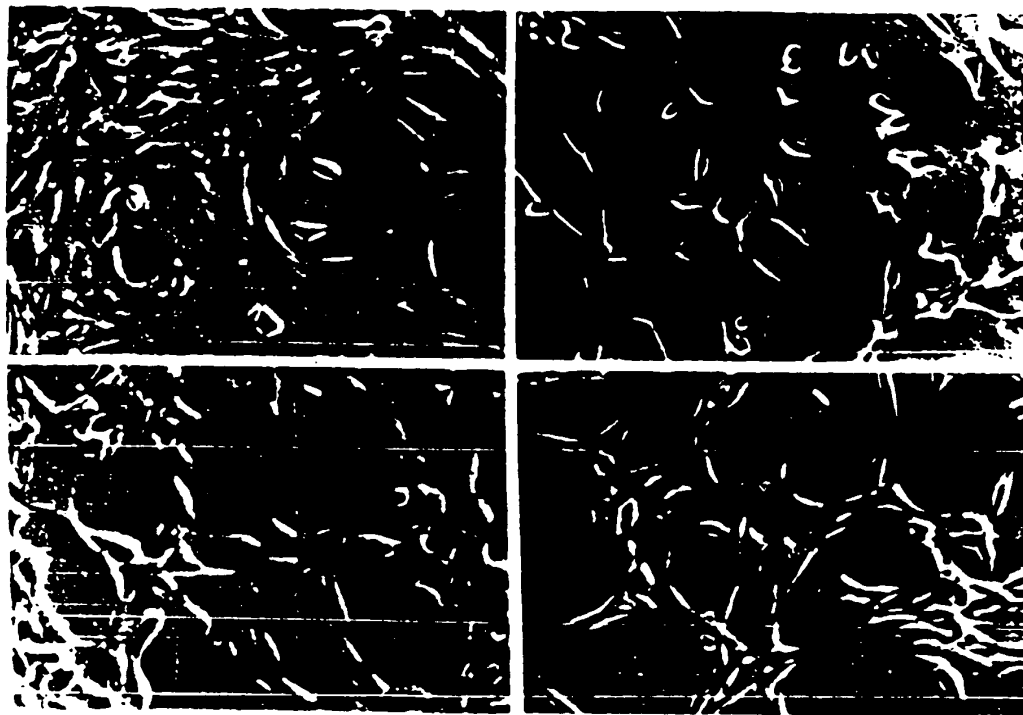
Figure 9. Morphology of NIH 3T3 cells transfected with wild-type or mutant HBGF-1 expression plasmids. The figure shows micrographs of the same NIH 3T3 cells analyzed by Western blot analysis in Fig. 8. A shows cells transfected with pSV2neo only and B–D show cells co-transfected with pSV2neo and expression vectors for wild-type (B and D) and mutant (C) HBGF-1. The cells shown in B correspond to those expressing relatively high levels of HBGF-1 (Fig. 8, lane 1), whereas those shown in D correspond to those expressing relatively little HBGF-1 (Fig. 8, lane 3).

sponse to PDGF Similarly, Severin···· et al. (40) used similar methods to generate a system where the mutant receptor could mediate an increase in c-fos expression in response to PDGF but not actin reorganization or mitogenesis.

The mitogenic deficiencies of the mutant HBGF-1 may be due to reduced biological stability in tissue culture medium, reduced binding to cell surface proteoglycans, an altered intracellular stability, and/or an altered affinity for an intracellular receptor or binding protein. It has been established that the presence of heparin protects HBGF-1 from thermal and proteolytic inactivation (28, 37). In addition, it has been shown that ¹²⁵I-labeled HBGF-1 is relatively insensitive to lysosomal degradation after receptor-mediated endocytosis (14). There is no obvious difference in the susceptibility of wild-type and mutant HBGF-1 to proteolytic cleavage by the conditioned media of NIH 3T3 cells cultured in the presence of 10% calf serum. However, the relative resistance of wild-type and mutant HBGF-1 to proteolytic modification in the presence of target cells or after receptor-mediated endocytosis has not been established. It is also possible that the mutant protein is more susceptible than the wild type to nonproteolytic inactivation. Further studies should reveal whether the altered activities of the mutant HBGF-1 are a consequence of its reduced apparent affinity for heparin.

In summary, the data presented here demonstrate that the various functions of HBGF-1 can be dissociated at the structural level. The observation that site-directed mutagenesis can be used to produce recombinant proteins with "normal" receptor-binding activity and reduced mitogenic activity indicates that similar methods could be used to produce potent antagonists of HBGF-1. More importantly, these results indicate that it may be possible through structure–function analysis and site-directed mutagenesis to generate mutants that retain certain (i.e., chemotactic, mitogenic, or heparin-binding) but not other biological functions characteristic of the wild-type protein. Finally, whereas the data presented on the receptor-binding and tyrosine kinase activation properties of the ρ132E mutant demonstrate that a lysine residue at this position is not critical for these functions, it is still possible that methylation of a lysine at this position could lead to reduced receptor-binding activity of HBGF-1 (19).

References

1. Baird, A., D. Schubert, N. Ling, and R. Guillemin. 1986. Receptor- and heparin-binding domains of basic fibroblast growth factor. Proc. Natl. Acad. Sci. USA. 83:2324–2328.
2. Bouche, G., N. Gas, H. Prats, V. Baldin, J.-P. Tauber, J. Teissié, and F. Amalric. 1987. Basic fibroblast growth factor enters the nucleolus and stimulates the transcription of ribosomal genes in ABAE cells undergoing G₀→G₁ transition. Proc. Natl. Acad. Sci. USA. 84:6770–6774.
3. Burgess, W. H., and T. Maciag. 1989. The heparin-binding (fibroblast)

4. Burgess, W. H., T. Mehlman, R. Friesel, W. V. Johnson, and T. Maciag. 1985. Multiple forms of endothelial cell growth factor: rapid isolation and biological and chemical characterization. J. Biol. Chem. 260:11389–11392.
5. ···· W. H., T. Mehlman, D. R. Marshak, B. A. Fraser, and T. Maciag. ··· Structural evidence that endothelial cell growth factor β is the precursor of an endothelial cell growth factor α and acidic fibroblast growth factor. Proc. Natl. Acad. Sci. USA. 83:7216–7220.
6. Burgess, W. H., C. A. Dionne, U. Kaplow, R. Mudd, R. Friesel, A. Zilberstein, J. Schlessinger, and M. Jaye. 1990. Characterization and cDNA cloning of phospholipase c-γ, a major substrate for heparin-binding growth factor 1 (acidic fibroblast growth factor)–activated tyrosine kinase. Mol. Cell. Biol. 10:4770–4777.
7. Casey, G., R. Smith, D. McGillivray, O. Peters, and C. Dickson. 1986. Characterization and chromosome assignment of the human homolog of int-2, a potential proto-oncogene. Mol. Cell. Biol. 6:502–510.
8. Coughlin, S. R., J. Barr, L. S. Cousens, L. J. Fretto, and L. T. Williams. 1988. Acidic and basic fibroblast growth factors stimulate tyrosine kinase activity in vivo. J. Biol. Chem. 263:988–993.
9. Dean, ···· P., and C. Basilico. 1987. Isolation of a rearranged human transforming gene following transfection of Kaposi sarcoma DNA. Proc. Natl. Acad. Sci. USA. 84:5660–5664.
10. Dickson, C., and G. Peters. 1987. Potential oncogene product related to growth factors. Nature (Lond.). 326:833–836.
11. Dickson, C., R. Smith, S. Brookes, and G. Peters. 1984. Tumorigenesis by mouse mammary tumor virus: proviral activation of a cellular gene in the common integration region int-2. Cell. 37:529–536.
12. Escobedo, J. A., and L. T. Williams. 1988. A PDGF receptor domain essential for mitogenesis but not for many other responses to PDGF. Nature (Lond.). 335:85–87.
13. Finch, P. W., J. S. Rubin, T. Miki, D. Ron, and S. Aaronson. 1989. Human KGF is FGF-related with properties of a paracrine effector of epithelial cell growth. Science (Wash. DC). 245:752–755.
14. Friesel, R., and T. Maciag. 1988. Internalization and degradation of heparin-binding growth factor-1 by endothelial cells. Biochem. Biophys. Res. Commun. 151:957–964.
15. Friesel, R., W. H. Burgess, and T. Maciag. 1989. Heparin-binding growth factor 1 stimulates tyrosine phosphorylation in NIH 3T3 cells. Mol. Cell. Biol. 9:1857–1865.
16. Friesel, R., W. H. Burgess, T. Mehlman, and T. Maciag. 1986. The characterization of the receptor for endothelial cell growth factor by covalent ligand attachment. J. Biol. Chem. 261:7581–7584.
17. Gay, C. G., and J. A. Winkles. 1990. Heparin-binding growth factor-1 stimulation of human endothelial cells induces platelet-derived growth factor A-chain gene expression. J. Biol. Chem. 265:3284–3292.
18. Gospodarowicz, D., G. Neufeld, and L. Schweigerer. 1987. Fibroblast growth factor: structural and biological properties. J. Cell. Physiol. 5:15–26.
19. Harper, J. W., and R. R. Lobb. 1988. Reductive methylation of lysine residues in acidic fibroblast growth factor: effect on mitogenic activity and heparin affinity. Biochemistry. 27:671–678.
20. Huang, S. S., and J. S. Huang. 1986. Association of bovine brain-derived growth factor receptor with protein tyrosine kinase activity. J. Biol. Chem. 261:9568–9571.
21. Jaye, M., R. Howk, W. Burgess, G. A. Ricca, I. Chiu, M. W. Ravera, S. J. O'Brien, W. S. Modi, T. Maciag, and W. N. Drohan. 1986. Human endothelial cell growth factor: cloning, nucleotide sequence, and chromosome localization. Science (Wash. DC). 233:541–545.
22. Jaye, M., W. H. Burgess, A. B. Shaw, and W. N. Drohan. 1987. Biological equivalence of natural bovine and recombinant human α-endothelial cell growth factor. J. Biol. Chem. 262:16612–16617.
23. Jaye, M., R. M. Lyall, R. Mudd, J. Schlessinger, and N. Sarver. 1988. Expression of acidic fibroblast growth factor cDNA confers growth advantage and tumorigenesis to Swiss 3T3 cells. EMBO (Eur. Mol. Biol. Organ.) J. 7:963–969.
24. Kan, M., D. DiSorbo, J. Hou, H. Hoshi, P. E. Mansson, and W. L. McKeehan. 1988. High and low affinity binding of heparin-binding growth factor to a 130-kDa receptor correlates with stimulation and inhibition of growth of a differentiated human hepatoma cell. J. Biol. Chem. 263:11306–11313.
25. Koda, T., A. Sasaki, S. Matsushima, and M. Kakinuma. 1987. A transforming gene, hst, found in NIH 3T3 cells transformed with DNA from stomach cancers and a colon cancer. Jpn. J. Cancer Res. 78:325–328.
26. Laemmli, U. K. 1970. Cleavage of structural proteins during the assembly of the head of bacteriophage T4. Nature (Lond.). 277:680–685.
27. Lobb, R. R. 1988. Clinical applications of heparin-binding growth factors. Eur. J. Clin. Invest. 18:321–336.
28. Lobb, R. R. 1988. Thrombin inactivates acidic fibroblast growth factor but not basic fibroblast growth factor. Biochemistry. 27:2572–2578.
29. Lobb, R. R., J. W. Harper, and J. W. Fett. 1986. Purification of heparin-binding growth factors. Anal. Biochem. 154:1–14.
30. Deleted in proof.
31. Margolis, B., S. G. Rhee, S. Felder, M. Mervic, R. Lyall, A. Levitzki, A. Ullrich, A. Zilberstein, and J. Schlessinger. 1989. EGF induces tyrosine phosphorylation of phospholipase C-γ: a potential mechanism for

Transforming Growth Factor α: Mutation of Aspartic Acid 47 and Leucine 48 Results in Different Biological Activities

ELIANE LAZAR,† SHINICHI WATANABE,* STEPHEN DALTON, AND
MICHAEL B. SPORN

Laboratory of Chemoprevention, National Cancer Institute, Bethesda, Maryland 20892

To study the relationship between the primary structure of transforming growth factor α (TGF-α) and some of its functional properties (competition with epidermal growth factor (EGF) for binding to the EGF receptor and induction of anchorage-independent growth), we introduced single amino acid mutations into the sequence for the fully processed, 50-amino-acid human TGF-α. The wild-type and mutant proteins were expressed in a vector by using a yeast α mating pheromone promoter. Mutations of two amino acids that are conserved in the family of the EGF-like peptides and are located in the carboxy-terminal part of TGF-α resulted in different biological effects. When aspartic acid 47 was mutated to alanine or asparagine, biological activity was retained; in contrast, substitutions of this residue with serine or glutamic acid generated mutants with reduced binding and colony-forming capacities. When leucine 48 was mutated to alanine, a complete loss of binding and colony-forming abilities resulted; mutation of leucine 48 to isoleucine or methionine resulted in very low activities. Our data suggest that these two adjacent conserved amino acids in positions 47 and 48 play different roles in defining the structure and/or biological activity of TGF-α and that the carboxy terminus of TGF-α is involved in interactions with cellular TGF-α receptors. The side chain of leucine 48 appears to be crucial either, indirectly in determining the biologically active conformation of TGF-α or directly in the molecular recognition of TGF-α by its receptor.

Transforming growth factor α (TGF-α) is a polypeptide of 50 amino acids. First isolated from a retrovirus-transformed mouse cell line (9), it has subsequently been found in human tumor cells (10, 29), in the early rat embryo (18), and recently in cell cultures from the pituitary gland (23). TGF-α appears to be closely related to epidermal growth factor (EGF) structurally and functionally (19, 20). The two peptides apparently bind to the same receptor, and both induce anchorage-independent growth of certain nontransformed cells, such as NRK cells, in the presence of TGF-β (1).

Comparison of amino acid sequences reveals about 35% homology among the EGF-like peptides (rat [27], mouse [25], and human [13] EGFs and rat [19] and human [12] TGF-αs). Some viral peptides (Shope fibroma growth factor [6], vaccinia growth factor [2], and myxoma growth factor [30]) also share homologies with the EGF-like peptides.

If TGF-α is involved in transformation, a TGF-α antagonist could be an important therapeutic tool in the treatment of certain types of malignancies. An understanding of the conformational and dynamic properties of the TGF-α molecule is basic to the design of an antagonist. A hypothetical antagonist would bind to the same receptor as TGF-α, but would not induce the series of proliferative and transforming events induced by TGF-α. To obtain such a molecule it is necessary to dissociate interactions responsible for binding from those involved in signal transduction. We decided to approach the problem by way of site-directed mutagenesis of a human sequence of TGF-α. In this report we describe our first series of mutations, which were carried out at residues Asp-47 and Leu-48, in the carboxy-terminal part of TGF-α; these two amino acids are highly conserved in the EGF-like family of peptides. We show that these two adjacent residues

play different roles in the structure and/or function of TGF-α.

MATERIALS AND METHODS

Cells. Normal rat kidney (NRK) cells were grown in Dulbecco modified Eagle medium containing 10% (vol/vol) calf serum.

TGF-α gene. The sequence of the 50-amino-acid human TGF-α was originally derived from a human TGF-α precursor cDNA (12). The coding sequence is preceded by an ATG methionine codon and followed by a TAA stop codon and is flanked by EcoRI restriction sites. This EcoRI fragment combines the 59-base-pair EcoRI-NcoI fragment from plasmid pTE5 (12) with the 111-base-pair NcoI-EcoRI fragment from plasmid pyTE2 (11). The resulting EcoRI fragment was inserted in M13mp18 for site-directed mutagenesis.

Synthesis and purification of oligonucleotides and oligonucleotide-directed mutagenesis. The synthesis and purification of 20- to 27-nucleotide oligonucleotides were carried out as described previously (31). The one or two nucleotides responsible for the mutation were located in the middle of the oligonucleotide. Mutagenesis was performed by published procedures (21, 33). The sequences of the mutant clones were verified by the method of Sanger et al. (25).

Yeast shuttle vector. The vector YEp70αT contains a yeast α-factor pheromone promoter and prepro sequence for the expression of TGF-α (15). The mutant TGF-α coding sequence was inserted in the EcoRI site of plasmid YEp70αT and expressed in the form of a fusion protein consisting of 92 amino acids from the prepro sequence of the yeast α factor attached to the amino terminus of TGF-α (28). The yeast cleaves the precursor and secretes TGF-α with 8 amino acids fused to it (4 are encoded by the prepro sequence of α-factor, and the other 4 are encoded by the DNA sequence added to insert of the TGF-α gene). The last of these residues is a methionine, which allows the cleavage of the secreted fusion

---
* Corresponding author.

† Present address: Unité d'Oncologie Moleculaire, IRSC, 94802 Villejuif Cedex, France.

protein by cyanogen bromide (CNBr) and the release of a mature TGF-α (50 amino acids) (see Results).

Yeast strain and transformation. The yeast *Saccharomyces cerevisiae* 20B-12 (*MATα trpl pep4-3*) (17) was obtained from the Yeast Genetics Stock Center, Berkeley, Calif. *S. cerevisiae* 20B-12 was grown in YEPD medium (1% yeast extract [Difco Laboratories], 2% Bacto-Peptone [Difco], 2% glucose). When the culture reached an optical density at 660 nm of 1, spheroplasts were prepared (14) for transformation. For each transformation we used 10 to 15 μg of purified plasmid DNA.

Partial purification of TGF-α mutants. At 3 days after transformation. five individual colonies of transformants were grown to saturation in YEPD medium. The amount of protein in the yeast medium was measured by the method of Bradford (3), and the amount of mutant TGF-α secreted in the yeast medium was determined by radioimmunoassay. The clones which secrete the highest amount of mutant TGF-α were used to grow a 1-liter culture in YNB-CAA medium (0.67% yeast nitrogen base, 20 g of glucose per liter, 10 g of Casamino Acids [Difco] per liter). After the culture reached saturation (optical density at 660 nm of 10 to 12) (48 h in an air shaker at 30°C), the yeast conditioned medium was dialyzed extensively against 1 M acetic acid in 3,000-molecular-weight cutoff dialysis tubing. Usually 250 ml of dialyzed culture was lyophilized, suspended in 10 ml of 70% formic acid, and treated with CNBr (molar excess of 500) for 20 h at room temperature. The CNBr was subsequently evaporated, and the samples were lyophilized. CNBr-treated samples were suspended in 1 ml of 1 M acetic acid, loaded on a Bio-gel P30 column (30 by 1.5 cm [Bio-Rad Laboratories]), and eluted with 1 M acetic acid. Fractions of 1 ml were collected. Aliquots were lyophilized, suspended in binding buffer (minimum essential medium containing 1 mg of bovine serum albumin per ml and 25 mM HEPES [*N*-2-hydroxy-ethylpiperazine-*N'*-2-ethanesulfonic acid; pH 7.4]), neutralized if necessary to pH 7.4, and tested in EGF-binding competition and soft-agar assays, as well in radioimmunoassay.

Radioimmunoassays. The amounts of TGF-α secreted in the yeast medium were determined by radioimmunoassay with the immunoglobulin G fraction of a polyclonal antibody, 34D, raised against recombinant human TGF-α (4), in 0.1 M Tris (pH 7.5)–0.15 M NaCl–2.5 mg of bovine serum albumin per ml. The amounts of partially purified TGF-α present in the P30 column fractions were measured by using the Biotope RIA kit with polyclonal antibody against human TGF-α (a gift from W. Hargreaves, Biotope), under denaturing conditions, as recommended by the supplier.

EGF binding competition assay and soft agar assay. Both EGF-binding competition and soft-agar assays have been described previously (1).

RESULTS

Rationale for mutations in the carboxyl terminus of TGF-α. Figure 1 shows the amino acid sequence of TGF-α in which the residues that are conserved among all the EGF-like peptides described thus far (EGF, TGF-α, and EGF-like viral proteins) are enclosed in bold circles. Among the 11 conserved amino acids, there are 6 Cys and 2 Gly residues, which presumably play essential roles in determining the overall conformation of the molecule. We concentrated on the two conserved amino acids in the carboxyl terminus. Asp-47 and Leu-48. The Asp in position 47 is conserved among the EGFs and TGF-α (human or murine), but not
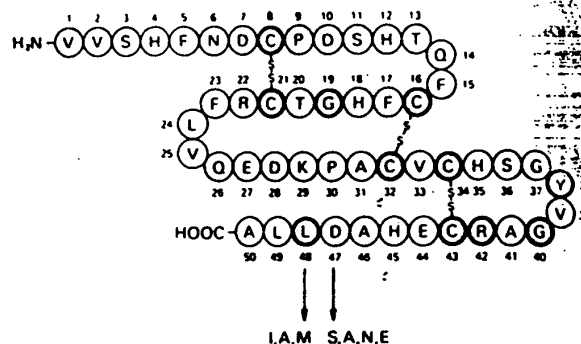
FIG. 1. Mutations in the carboxy terminus of human TGF-α. The amino acids conserved in all the family of EGF-like growth factors (human and murine EGFs and TGFs, as well as the gene products of the vaccinia virus [vaccinia growth factor], the Shope fibroma virus [Shope fibroma growth factor], and the myxoma virus [myxoma growth factor]) are enclosed in bold circles. The mutations of amino acids at positions 47 and 48 are indicated. Symbols: A, Ala; C, Cys; D, Asp; E, Glu; F, Phe; G, Gly; H, His; I, Ile; K, Lys; L, Leu; M, Met; N, Asn; P, Pro; Q, Gln; R, Arg; S, Ser; T, Thr; V, Val; W, Trp; Y, Tyr.

among the EGF-like viral proteins (vaccinia growth factor, Shope fibroma growth factor, or myxoma growth factor), whereas Leu 48 is conserved among all the EGF-like peptides so far described. In both mouse and human EGF, the two corresponding residues (Asp-46 and Leu-47) are located near the surface of the protein (8, 22, 22a). We designed a series of mutations in these two positions.

Asp-47 has been mutated to Glu, Asn, Ser, and Ala. Glu was chosen because it has the same charge as and a larger size than Asp: Asn has a similar side-chain structure, but is uncharged: Ser is smaller but still polar; Ala is smaller and nonpolar.

Leu 48 has been mutated to Ile and Met, which are both large, nonpolar residues like Leu, and to Ala, which is nonpolar but smaller. We introduced the chosen mutations by site-directed mutagenesis of the cloned human TGF-α gene, using synthetic oligonucleotides.

Construction of the yeast α mating pheromone–human TGF-α plasmid. The TGF-α expression vector pyTE1 (Fig. 2) was constructed by using plasmid YEp70αT (15) which contains the 2μm origin of replication and yeast *TRP1* gene for its replication and selective maintenance, respectively. YEp70αT also contains the yeast α-factor promoter, the α-factor prepro sequence coding for 89 amino acids, and the sequence for 3 amino acids resulting from the introduction of *Xba*I and *Eco*RI sites. The human mature TGF-α sequence (12) is contained in a 170-base-pair *Eco*RI fragment which includes an ATG (Met) codon preceding the sequence of TGF-α and a TAA (stop) codon followed by 8 nucleotides. This TGF-α sequence was inserted in the unique *Eco*RI site of YEp70αT. Clones with the proper orientation were selected, and DNA was isolated for yeast transformation.

Measurement of TGF-α secreted by *S. cerevisiae*. The amount of total proteins secreted into the yeast culture was 10 ± 1 μg/ml for wild-type as well as mutant TGF-α as determined by the method of Bradford (3). Before further purification was attempted, we wanted to determine whether the mutated TGF-α proteins were being secreted by the yeast. The low pH of the yeast medium, as well as the acidic proteins secreted in the yeast culture, precluded biological assay of secreted mutants. Therefore, immunological meth-

M   V   V .... A   Stop
ATGGTGGTG................. GCCTAA
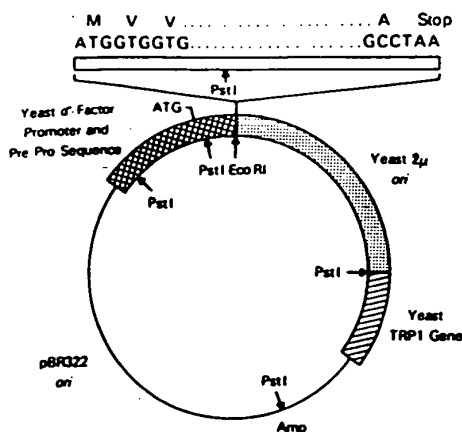
FIG. 2. Structure of the *S. cerevisiae* 8.2-kilobase shuttle vector pyTE1. The secretion of the TGF-α gene is under the transcriptional control of the yeast α-factor promoter and prepro sequence (▨). The yeast 2μm origin of replication (▨) and the selective yeast *TRP1* gene (▨) are indicated. The TGF-α gene, preceded by an initiation (ATG) codon and followed by a stop (TAA) codon, is inserted in the *Eco*RI site. Details are given in Materials and Methods and in Results.

FIG. 3. Purification of yeast-secreted wild-type TGF-α. The purification procedure is described in Materials and Methods and in Results. Aliquots of every other fraction of the Bio-Gel P30 column were tested for their abilities to compete with ^{125}I-EGF for binding to the EGF receptor (Δ) and to induce colony formation (>62 μm) on NRK cells in soft agar in the presence of TGF-β (1 ng/ml) (●). The A_{280} profile of the proteins was determined (——).

ods were used. Wild-type and mutant TGF-α's were secreted at a level of 100 to 200 ng/ml and 10 to 500 ng/ml, respectively (as determined by radioimmunoassay with polyclonal antibody 34D). We thus estimate that the percentage of TGF-α secreted in the yeast culture is at least 1% of the total protein secreted. We cannot yet assess whether the variations in the levels of secretion of different mutant TGF-α proteins are real or whether one single-amino-acid substitution drastically affects the recognition by the antibody. The latter hypothesis is the more likely, since the use of an ther polyclonal antibody (Biotope) under denaturing conditions enabled us to detect certain TGF-α mutants (such as [Ala 47]-TGF-α, in which the amino acid in position 47 of human TGF-α is mutated to an alanine) that were poorly detected by 34D, under nondenaturing as well as denaturing conditions. After the amount of TGF-α mutant proteins was estimated, the medium was extensively dialyzed against 1 M acetic acid and lyophilized as described in Materials and Methods.

Partial purification of yeast-secreted TGF-α. Although the yeast shuttle vector was constructed in such a way as to secrete TGF-α with 8 amino acids fused to the N terminus, it was often observed that a significant fraction of the secreted TGF-α was in a higher-molecular-weight fragment corresponding to the size expected from an uncleaved (unprocessed) 92-amino-acid fusion protein. Since a Met had been introduced at the N terminus of TGF-α and since TGF-α contains no Met in its sequence, CNBr treatment could be used to cleave either of these 8- or 92-amino-acid N-terminal peptides and release the complete 50-amino-acid TGF-α. Indeed, CNBr treatment of yeast-secreted proteins resulted in the conversion of high-molecular-weight TGF-α into the 6,000-molecular-weight species, as revealed by Western immunoblot (data not shown).

CNBr-cleaved samples (see Materials and Methods) were purified on a Bio-Gel P30 column. Figure 3 shows the elution profile of the proteins, as well as the results of a radioreceptor assay and a soft-agar assay performed on aliquots of the column fractions. The A_{280} profile shows two major peaks of

eluted proteins, one corresponding to the void volume and the other one to proteins of molecular weight <3,000. Aliquots of the column fractions were tested for their ability to compete with ^{125}I-EGF for binding to the receptor. The fractions that were the most active in this assay were located between the two major protein peaks, in an area where relatively few proteins eluted. Although some activity was found in the first protein peak (void volume), this was considerably reduced on treatment with stronger CNBr (data not shown).

Aliquots of each fraction were also tested for their ability to induce anchorage-independent growth of NRK cells in soft agar in the presence of TGF-β (1 ng/ml). The receptor binding and colony-forming activity superimposed almost exactly (Fig. 3). Analysis by polyacrylamide gel electrophoresis with silver staining, as well as by Western blot, of the column fractions shows that our purification procedure (CNBr cleavage followed by P30 sizing column) eliminates high-molecular-weight proteins (data not shown). Since pure TGF-α migrates in a broad band on sodium dodecyl sulfate-polyacrylamide gel electrophoresis (32), this technique cannot be used for proper assessment of the degree of separation of TGF-α from low-molecular-weight contaminating proteins. Nevertheless, within our detection levels the amounts of TGF-α present in the column fractions (detected by radioimmunoassay using the antibody from Biotope) correlated with the amounts observed on sodium dodecyl sulfate-polyacrylamide gel electrophoresis (data not shown).

Comparison of binding and colony-forming activity of TGF-α partially purified from yeast media. It was important to show that wild-type TGF-α secreted from *S. cerevisiae* had the expected biological properties and that its activity in soft-agar and radioreceptor assays was equivalent. For these assays, the amount of EGF-competing activity present in the most active fraction of the P30 column of wild-type TGF-α was measured in terms of EGF equivalents. The dilution curve had a slope that was parallel to that of the EGF standard. This value was also used to measure the colony-forming activity of the partially purified wild-type TGF-α (with EGF as a standard in the assay). The colony-forming activity of the partially purified wild-type TGF-α corre-
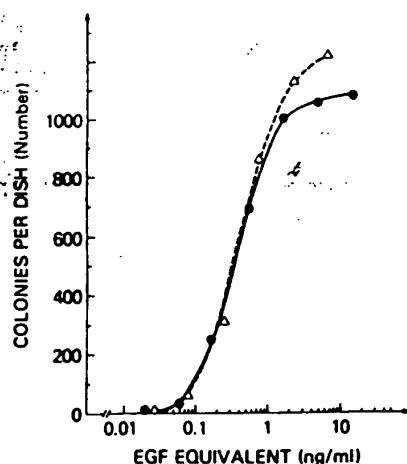
FIG. 4. Correlation between the activities in the binding and colony-forming assay for the partially purified wild-type TGF-α secreted by *S. cerevisiae*. The activity in the radioreceptor assay of the peak fraction from the P30 column was determined in EGF equivalent concentration. The value obtained was used for the soft-agar assay. Colonies of >62 μm (Δ) and the EGF standard (●) are shown.

TABLE 1. Biological and biochemical activities of mutant TGF-α proteins secreted by *S. cerevisiae* and partially purified

Insert in the yeast expression vector	EGF equivalence (ng/ml) in:		Amt of TGF-α (ng/ml) in radioimmunoassay
	Radioreceptor assay	Soft-agar assay	
Wild-type TGF-α	700	700	2,000
	400	300	ND[a]
None	0	0	0
[Ala-47]-TGF-α	100	44	220
	66	48	ND
[Asn-47]-TGF-α	80	72	180
	75	72	525
[Glu-47]-TGF-α	3	3	42
[Ser-47]-TGF-α	10	4	60
[Ala-48]-TGF-α	0	0	16
	0	0	220
[Ile-48]-TGF-α	4	12	470
	2	7	490
[Met-48]-TGF-α	2	8	453
	0.5	2	420

[a] ND, Not determined.

sponded exactly to that of EGF (Fig. 4). Thus, we have partially purified a wild-type 50-amino-acid TGF-α showing the expected binding and colony-forming activities, which provides a reference substance for mutant TGF-αs that might show a dissociation of binding and colony-forming abilities.

Biological and biochemical activities of the partially purified TGF-α mutant proteins. Mutated TGF-αs were expressed by using the yeast system and partially purified on Bio-Gel P30 columns as described in Materials and Methods. Mutant TGF-αs were usually obtained from two different clones of yeast transformants. The CNBr-cleaved samples were purified through different Bio-Gel P30 columns for each mutant protein to avoid any possible contamination from one peptide to another. The purification profiles observed with the mutant TGF-αs were similar to those obtained for the wild-type TGF-α. Aliquots of the P30 column fractions were tested in radioreceptor and soft-agar assays. For all mutant proteins, the highest activity in both assays was always found in the same fraction of the Bio-Gel P30 column effluent (peak fraction). Extensive purification of a series of mutant proteins for screening purposes is not practical. Therefore, we needed a quantitation system that would allow us to compare mutant proteins with each other. Thus, the amount of TGF-α present in the peak fraction was estimated by radioimmunoassay with an antiserum to native TGF-α (obtained from W. Hargreaves), under denaturing conditions, as described in Materials and Methods. All values given in Table 1 were obtained from the peak fraction.

The controls done with the wild-type TGF-α showed (Fig. 4; Table 1) that binding and transforming activity were equivalent. The yeast vector without a TGF-α insert did not secrete any EGF-like proteins, as determined by both radioreceptor and soft-agar assay.

Two types of results were obtained upon assay of mutant proteins having different amino acid substitutions at Asp-47. In both [Ala-47]-TGF-α and [Asn-47]-TGF-α, binding ability was retained. Soft-agar and radioreceptor activities correlated for [Asn-47]-TGF-α; there was a lower value for

colony-forming activity than for EGF-binding competition for [Ala-47]-TGF-α. [Ser-47]-TGF-α and [Glu-47]-TGF-α appeared to have lower activities in both assays than either wild-type TGF-α or [Ala-47]-TGF-α and [Asn-47]-TGF-α. These results indicate that neither the carboxyl charge nor the polarity of Asp-47 is essential for biological activity.

The effects of mutation of Leu-48, one of the 11 amino acids perfectly conserved among all the EGFs, TGF-αs, and viral EGF-like proteins, are dramatic. [Ala-48]-TGF-α totally lacked binding and colony-forming activity. [Ile-48]-TGF-α and [Met-48]-TGF-α had very little biological activity compared with wild-type TGF-α. Another substitution, [Met-48]-TGF-α, resulted in a truncated mutant lacking the last 2 amino acids and having a substitution of Leu to homoserine at position 48 following treatment with CNBr. Alternatively, if [Met-48]-TGF-α was not treated with CNBr, fusion proteins of TGF-α (mutated to Met in position 48) with 8 or 92 amino acids attached at the N terminus were obtained. Very low activities in binding and soft-agar assays were found for these mutants, whether or not they were cleaved with CNBr. Experiments on EGF and TGF-α have shown that an N-terminal extension does not markedly modify EGF-binding activity (12, 26). Therefore, the loss of activity obtained with [Met-48]-TGF-α that has not been CNBr treated was probably due to the mutation itself and not to the N-terminally extended fusion protein. We do not know whether the loss of activity observed with the TGF-α shortened to 48 amino acids and having a substitution of Leu-48 to homoserine is due only to the mutation or also to the lack of the last 2 amino acids.

The data obtained by radioimmunoassay on the partially purified wild-type and mutant TGF-α show that the amount of TGF-α detected was always higher than the amount determined by measurement of biological activity. This may be due to the presence in the fraction of a certain percentage of incorrectly folded TGF-α that might be recognized in a

radioimmunoassay under denaturing conditions but would not be biologically active. None of the mutant proteins seemed to be present in amounts equivalent to those observed for wild-type TGF-α in the partially purified fractions (whether radioimmunoassay, radioreceptor, or soft-agar assay was used for quantitation). It is not clear whether consistently less TGF-α was produced by the mutant constructs than by the wild type or whether the secreted mutant proteins were simply less well recognized by the antibody. Because of these uncertainties, the biological activities of the different mutant proteins cannot be accurately related to a known amount of mutant TGF-α protein. Even though radioimmunoassay should be used with caution for a quantitative evaluation of mutant TGF-α proteins, a positive reaction demonstrates that immunoreactive TGF-α was present in the P30 peak fraction for each mutant. Therefore, the fact that one of the mutant proteins ([Ala-48]-TGF-α) is biologically inactive can be attributed to the mutation itself, and not to the lack of production of the mutant protein by the yeast or its loss through purification. However, if the mutant proteins are in fact as immunoreactive as the wild type, then [Ala-47]-TGF-α and [Asn-47]-TGF-α are as active as wild-type TGF-α and [Glu-47]-TGF-α and [Ser-47]-TGF-α are less active; in contrast, [Ile-48]-TGF-α and [Met-48]-TGF-α are almost inactive. The differences between mutation of Asp-47 and Leu-48 would then be even more striking.

DISCUSSION

TGF-α shows sequence homologies with EGF, and both growth factors share the same cellular receptors (20). Even though EGF was discovered 25 years ago (7) and its properties have been extensively studied over the years (5), the binding site of EGF to its receptor has still not been determined, and the relationship between structure and function of EGF/TGF-α is still to be discovered. Particularly, we do not know whether binding to the receptor and signal transduction occur through one or more domains of the molecule or through which amino acids. We approached the question by performing site-directed mutagenesis of TGF-α and focused our attention on two adjacent amino acids, Asp-47 and Leu-48, located in the carboxy terminus and highly conserved in the EGF-like family of peptides. Unexpectedly, these two amino acids showed very different sensitivities to mutation and particularly to a substitution to Ala: [Ala-47]-TGF-α retained binding and colony-forming activities, whereas [Ala-48]-TGF-α completely lost both activities. These data show that Asp-47 and Leu-48 play very different roles in defining the structure and/or the activity of TGF-α. The other mutations performed on Asp-47 were substitutions to Asn, Ser, and Glu. [Asn-47]-TGF-α, like [Ala-47]-TGF-α, was active in binding and induction of colony formation, but [Ser-47]-TGF-α and [Glu-47]-TGF-α showed weaker growth factor activities. These results indicate that neither the carboxyl charge nor the polarity of Asp-47 is essential for biological activity. Interestingly, two of the EGF-like viral proteins, myxoma growth factor and Shope fibroma growth factor (6, 30), have Asn instead of Asp in position 47; we have shown that [Asn-47]-TGF-α retains biological activity.

Substitution of Leu-48 to Met and Ile led to mutant proteins with very low activities, whereas substitution to Ala led to complete loss of activity. We did not expect that a mutation of Leu to Ile (which have similar sizes and polarities) would cause such a strong effect. Thus, Leu-48, which is conserved perfectly among all the EGF-like peptides,

seems to be essential, through its exact geometry, f r the biological activity of TGF-α.

The mutant proteins tested so far, when active, showed parallel behaviors in binding and colony formation. Some mutant proteins lost all activities, and we assume that the binding capacity has been lost. We have not been able to dissociate the binding and colony-forming abilities by using any of the present series of mutant proteins, and it is necessary to screen more of them in search of an antagonist of TGF-α.

Results relating to the biological activity of EGF show that derivatives of mouse EGF and human EGF (EGF 1-47) lacking the carboxy-terminal 6 amino acids as a result of enzymatic digestion are less potent than the intact molecule in mitogenic stimulation of fibroblasts, but retain full biological activity in in vivo assays (inhibition of gastric acid secretion) (16). On the other hand, naturally occurring truncated forms of rat EGF, which lack the carboxy-terminal 5 amino acids (rEGF 2-48) are as potent as mouse EGF (mEGF 1-53) in receptor-binding and mitogenic assays (27). We do not know whether the discrepancies observed are due to the origin of the molecule (artificial or natural) or to the type of bioassay used. In any event, all of these EGF-related molecules, which are shorter than mouse or human EGF, still retain Leu-47. We have shown that in TGF-α, the corresponding residue, Leu-48, is critical for the biological activity.

Recent data on the three-dimensional structure of mouse EGF obtained by nuclear magnetic resonance show that even though Asp-46 and Leu-47 (Asp-47 and Leu-48 in TGF-α) are both solvent accessible (8, 22, 22a), their side chains point in opposite directions in the beta-sheet structure. Therefore, the role of these adjacent amino acids in the structure and, consequently, the function of EGF might be very different. Our data show that the amino acids Asp-47 and Leu-48 of TGF-α are not equally important for the biological activity of TGF-α, despite their conservation among the EGF-like peptides. From the dramatic loss in biological activity which is characteristic of mutation of Leu-48, we also suggest that this residue is involved in binding to the cellular receptors either by direct interaction with the receptor or by providing the proper conformation to the molecule.

LITERATURE CITED

1. Anzano, M. A., A. B. Roberts, J. M. Smith, M. B. Sporn, and J. E. De Larco. 1983. Sarcoma growth factor from conditioned medium of virally transformed cells is composed of both type α and type β transforming growth factors. Proc. Natl. Acad. Sci. USA 80:6264–6268.
2. Blomquist, M. C., L. T. Hunt, and W. C. Barker. 1984. Vaccinia virus 19-kilodalton protein: relationship to several mammalian proteins, including two growth factors. Proc. Natl. Acad. Sci. USA 81:7363–7367.
3. Bradford, M. M. 1976. A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. Anal. Biochem. 72:248–254.

4. Bringman, T. S., P. B. Lindquist, and R. Derynck. 1987. Different transforming growth factor-α species are derived from a glycosylated and palmitoylated transmembrane precursor. Cell 48:429–440.

5. Carpenter, G., and S. Cohen. 1979. Epidermal growth factor. Annu. Rev. Biochem. 48:193–216.

6. Chang, W., C. Upton, S. Hu, A. F. Purchio, and G. McFadden. 1987. The genome of Shope fibroma virus, a tumorigenic poxvirus, contains a growth factor gene with sequence similarity to those encoding epidermal growth factor and transforming growth factor α. Mol. Cell. Biol. 7:535–540.

7. Cohen, S. 1962. Isolation of a mouse submaxillary gland protein accelerating incisor eruption and eyelid opening in the new-born animal. J. Biol. Chem. 237:1555–1562.

8. Cooke, R. M., A. J. Wilkinson, M. Baron, A. Pastore, M. J. Tappin, I. D. Campbell, H. Gregory, and B. Sheard. 1987. The solution structure of human epidermal growth factor. Nature (London) 327:339–341.

9. De Larco, J. E., and G. J. Todaro. 1978. Growth factors from murine sarcoma virus-transformed cells. Proc. Natl. Acad. Sci. USA 75:4001–4005.

10. Derynck, R., D. V. Goeddel, A. Ullrich, J. U. Gutterman, R. D. Williams, T. S. Bringman, and W. H. Berger. 1987. Synthesis of messenger RNAs for transforming growth factors α and β and the epidermal growth factor receptor by human tumors. Cancer Res. 47:707–712.

11. Derynck, R., A. B. Roberts, D. H. Eaton, M. E. Winkler, and D. V. Goeddel. 1985. Human transforming growth factor-α: precursor sequence, gene structure, and heterologous expression. Cancer Cells 3:79–86.

12. Derynck, R., A. B. Roberts, M. E. Winkler, E. Y. Chen, and D. V. Goeddel. 1984. Human transforming growth factor-α: Precursor structure and expression in E. coli. Cell 38:287–297.

13. Gregory, H. 1975. Isolation and structure of urogastrone and its relationship to epidermal growth factor. Nature (London) 257:325–327.

14. Hinnen, A., J. B. Hicks, and G. R. Fink. 1978. Transformation of yeast. Proc. Natl. Acad. Sci. USA 75:1929–1933.

15. Hitzeman, R. A., C. N. Chang, M. Matteucci, L. J. Perry, W. J. Kohr, J. J. Wulf, J. R. Swartz, C. Y. Chen, and A. Singh. 1986. Construction of expression vectors for secretion of human interferons by yeast. Methods Enzymol. 119:424–433.

16. Hollenberg, M. D., and H. Gregory. 1980. Epidermal growth factor—urogastrone: biological activity and receptor binding of derivatives. Mol. Pharmacol. 17:314–320.

17. Jones, E. 1976. Proteinase mutants of Saccharomyces cerevisiae. Genetics 85:23–30.

18. Lee, D. C., R. Rochford, G. J. Todaro, and L. P. Villarreal. 1985. Developmental expression of rat transforming growth factor α mRNA. Mol. Cell. Biol. 5:3644–3646.

19. Marquardt, H., M. W. Hunkapiller, L. E. Hood, and G. J. Todaro. 1984. Rat transforming growth factor type 1: structure and relation to epidermal growth factor. Science 223:1079–1082.

20. Massague, J. 1983. Epidermal growth factor-like transforming growth factor. II. Interaction with epidermal growth factor

receptors in human placenta membranes and A431 cells. J. Biol. Chem. 258:13614–13620.

21. Messing, J. 1983. New M13 vectors for cloning. Methods Enzymol. 101:20–78.

22. Montelione, G. T., K. Wuthrich, E. C. Nice, A. W. Burgess, and H. A. Scheraga. 1986. Identification of two anti-parallel beta sheet conformations in the solution structure of murine epidermal growth factor by proton magnetic resonance. Proc. Natl. Acad. Sci. USA 83:8594–8598.

22a.Montelione, G. J., K. Wuthrich, E. C. Nice, A. W. Burgess, and H. A. Scheraga. 1987. Solution structure of murine epidermal growth factor: determination of the polypeptide backbone chain-fold by nuclear magnetic resonance and distance geometry. Proc. Natl. Acad. Sci. USA 84:5226–5230.

23. Samsoondar, J., M. S. Kobrin, and J. E. Kudlow. 1986. α-Transforming growth factor secreted by untransformed bovine anterior pituitary cells in culture. J. Biol. Chem. 261:14408–14413.

24. Sanger, F., S. Nicklen, and A. R. Coulson. 1977. DNA sequencing with chain-terminating inhibitors. Proc. Natl. Acad. Sci. USA 74:5463–5467.

25. Savage, C. R., Jr., J. H. Hash, and S. Cohen. 1973. Epidermal growth factor: location of disulfide bonds. J. Biol. Chem. 248:7669–7672.

26. Shechter, Y., J. Schlessinger, S. Jacobs, K. J. Chang, and P. Cuatrecasas. 1978. Fluorescent labeling of hormone receptors in viable cells: preparation and properties of highly fluorescent derivatives of epidermal growth factor and insulin. Proc. Natl. Acad. Sci. USA 75:2135–2139.

27. Simpson, R. J., J. A. Smith, R. L. Moritz, M. J. O'Hare, P. S. Rudland, J. R. Morrison, C. J. Lloyd, B. Grego, A. W. Burgess, and E. C. Nice. 1985. Rat epidermal growth factor: complete amino acid sequence. Eur. J. Biochem. 153:629–637.

28. Singh, A., J. M. Lugovoy, W. J. Kohr, and L. J. Perry. 1984. Synthesis, secretion and processing of α-factor-interferon fusion proteins in yeast. Nucleic Acids Res. 12:8927–8938.

29. Todaro, G. J., C. Fryling, and J. E. DeLarco. 1980. Transforming growth factors produced by certain human tumors: polypeptides that interact with epidermal growth factor receptors. Proc. Natl. Acad. Sci. USA 77:5258–5262.

30. Upton, C., J. L. Macen, and G. McFadden. 1987. Mapping and sequencing of a gene from myxoma virus that is related to those encoding epidermal growth factor and transforming growth factor α. J. Virol. 61:1271–1275.

31. Watanabe, S., E. Lazar, and M. B. Sporn. 1987. Transformation of normal rat kidney (NRK) cells by an infectious retrovirus carrying a synthetic rat type α transforming growth factor gene. Proc. Natl. Acad. Sci. USA 84:1258–1262.

32. Winkler, M. E., T. Bringman, and B. J. Marks. 1986. The purification of fully active recombinant transforming growth factor-α produced in Escherichia coli. J. Biol. Chem. 261:13838–13843.

33. Zoller, M. J., and M. Smith. 1983. Oligonucleotide-directed mutagenesis of DNA fragments cloned into M13 vectors. Methods Enzymol. 101:468–500.

AFFINITY CHROMATOGRAPHY APPLICATIONS INDEX

The table is intended as a general guide to selecting an affinity chromatography medium. Where a medium is suggested for a very specific application, literature references are given; those references should be consulted for appropriate chromatographic conditions and limitations of the method.

References are also given to support more general guidelines, such as ovomucoid-agarose for lectins. Where specific references to your particular purification problem do not exist, the general suggestions may help you select an appropriate medium. Such applications must be viewed as investigatory, however; you should expect to have to do some or many trials to arrive at the best combination of medium and conditions.

For Purification/Binding of	Suggested ligand
adenosine monophosphate	acriflavine[2]
albumin	cholic acid[3]
amino acids, D and L	albumin[86]
antibodies	protein A; protein G; anti-Ig; serum proteins
avidin-labeled compounds	biotin
bilirubin	albumin[87]
biotin and biotin-labeled compounds	avidin; streptavidin
cells	lectins on macroagarose (agarose macrobeads)
halophilic bacteria	ω-aminohexyl[118]
HeLa cells	lens culinaris lectin[84]
human erythrocytes	anti-rabbit IgG[88]
mouse thymocytes	protein A[37]
chorionic gonadotropin	concanavalin A[8]
coagulation factors	heparin[13]
coagulation factor II	sulfated dextran[14]; heparin[13]
plasminogen	L-lysine[82]
plasminogen activators	p-aminobenzamidine[82]
thrombin	m-aminobenzamidine[88]; p-aminobenzamidine[83]
dihydrofolate reductase	Cibacron blue 3GA[116]
endotoxins (pyrogens)	histamine[15]; polymyxin B[16]
enzymes	amino acids; dyes; nucleotides/cofactors
N-acetylhexosaminidase	concanavalin A[8]
adenosine deaminase	ω-aminohexyl[122]
adenosine kinase	adenosine 5'-monophosphate[78]
alcohol dehydrogenase	Cibacron blue 3GA[116]; adenosine 5'-monophosphate[78]
alkaline phosphatase	L-histidyldiazobenzylphosphonic acid[20]
aminoacyl tRNA-synthetase	ω-aminohexyl[119]
aminopeptidase	ω-aminohexyl[133]
anhydrochymotrypsin	trypsin inhibitor[26]
anthranilate synthetase	L-tryptophan[85]
carbohydrate-metabolizing enzymes	carbohydrates
carbonic anhydrase	p-aminomethylbenzenesulfonamide[22]
carboxypeptidase A	D-tryptophan[84]
casein kinase	α-casein[83]; heparin[82]
cholesterol oxidase	cholesteryl hemisuccinate[23]
α-chymotrypsin	4-phenylbutylamine[24]
cytochrome oxidase	cytochrome c[25]
DAHP synthetase, Tyr-sensitive	L-tyrosine[86]
dehydrogenases	nucleotides; NAD; NADP; Cibacron blue 3GA[116]
	reactive red 120[116]
detoxifying enzymes	glutathione
dihydrofolate reductase	methotrexate[30] ; Cibacron blue 3GA[116]
dihydroneopterin triphosphate synthetase	guanosine 5'-triphosphate[74]
enterokinase	p-aminobenzamidine[43]
α-galactosidase	D(+)melibiose[104]
β-galactosidase	p-aminobenzyl 1-thio-β-D-galactopyranoside[31]
	p-aminophenyl β-D-thiogalactopyranoside[31]
	p-aminobenzyl 1-thio-β-D-galactopyranoside[31]
galactosyltransferase	α-lactalbumin[32]; N-acetyl-D-glucosamine[88]
glucuronidase	saccharolactone[109]
glutamate dehydrogenase	ω-aminopentyl
glutamate synthetase	Cibacron blue 3GA[8]
glutathione S-transferase	cholic acid[34]
glyceraldehyde 3-phosphate dehydrogenase	reactive blue 72; pentyl
glyceraldehyde dehydrogenase	adenosine 5'-monophosphate[77]
glycogen phosphorylase b	ω-aminohexyl[121] ; butyl[114]
glycogen synthetase	ω-aminobutyl[121]; uridine 5'-diphosphate[71]
hexokinase	N-acetyl-D-glucosamine[104]
histone kinase	heparin[102]
kallikrein	trypsin inhibitor[21]
kinases	nucleotides; ATP; Cibacron blue 3GA[16]
lactic dehydrogenase	β-nicotinamide adenine dinucleotide[72]

Left margin fragments:
ng the buf-
lumn at a
used to elute
ing effects
required for
order to have
capacity

US $
25 ml 43.45
100 ml 120.60
200 ml 201.00

fer 96

ve com-
most listings

we offer a com-

Press,

AMINO ACID RESINS

Immobilized amino acids are very versatile affinity media. Their use in the isolation of proteins and enzymes is well established. Recently, there has been increasing interest in the use of these resins for serum protein separation.

General References:
Deutsch, D.G. and Mertz, E.T., Proc. Fed. Amer. Soc. Exp. Biol., 29, 647 (1970).
Deutsch, D.G. and Mertz, E.T., Science, 170, 1095 (1970).
Vverto, M. and Vaheki, A., Biochem. J., 183, 331 (1979).

PRODUCT NUMBER		US $
D-ALANINE A 3435	Matrix: Cross-linked 4% beaded agarose Activation: cyanogen bromide Attachment: amino Spacer: 1 atom Ligand immobilized: 2-8 μmoles per ml Form: Suspension in 2.0 M NaCl containing 0.02% thimerosal	1 ml 13.30 5 ml 43.20 10 ml 71.45
L-ALANINE A 3519	Matrix: 4% beaded agarose Activation: cyanogen bromide Attachment: amino Spacer: 1 atom Ligand immobilized: 2-7 μmoles per ml Form: Suspension in 2.0 M NaCl containing 0.02% thimerosal	1 ml 13.30 5 ml 43.20 10 ml 71.45 25 ml 142.20
L-ARGININE A 1018	Matrix: 4% beaded agarose Activation: cyanogen bromide Attachment: amino Spacer: 1 atom Ligand immobilized: 5-10 μmoles per ml Form: Suspension in 2.0 M NaCl containing 0.02% thimerosal	1 ml 13.90 5 ml 44.70 10 ml 74.25 25 ml 148.20
A 8405	Matrix: Cross-linked 4% beaded agarose Activation: epoxy Attachment: amino Spacer: 12 atoms Ligand immobilized: 1-3 μmoles per ml Form: Lyophilized powder stabilized with lactose Swelling: 1 g swells to approx. 12 ml	1 ml 10.80 5 ml 34.95 10 ml 57.80
L-ASPARAGINE A 1143	Matrix: 4% beaded agarose Activation: cyanogen bromide Attachment: amino Spacer: 1 atom Ligand immobilized: 5-10 μmoles per ml Form: Suspension in 2.0 M NaCl containing 0.02% thimerosal	1 ml 13.30 5 ml 43.20 10 ml 71.45 25 ml 142.20
A 2047	Matrix: Cross-linked 4% beaded agarose Activation: epoxy Attachment: amino Spacer: 12 atoms Ligand immobilized: 1-2 μmoles per ml Form: Suspension in 2.0 M NaCl containing 0.02% thimerosal	1 ml 13.30 5 ml 43.20 10 ml 71.45 25 ml 142.20
L-ASPARTIC ACID A 3394	Matrix: 4% beaded agarose Activation: cyanogen bromide Attachment: amino Spacer: 1 atom Ligand immobilized: 5-10 μmoles per ml Form: Suspension in 2.0 M NaCl containing 0.02% thimerosal	1 ml 21.20 5 ml 68.50 10 ml 113.80 25 ml 227.15

PRODUCT NUMBER		US $
L-CYSTEINE C 7896	Matrix: 4% beaded agarose Activation: cyanogen bromide Attachment: amino Spacer: 1 atom Ligand immobilized: 0.5-1.0 μmole per ml Form: Suspension in 0.5 M NaCl, 0.01 M citrate, pH 4.5, containing 0.01% thimerosal	1 ml 15.60 5 ml 49.70 10 ml 82.50 25 ml 164.10
L-CYSTINE C 9266	Matrix: agarose	Inquire
L-GLUTAMIC ACID G 2759	Matrix: 4% beaded agarose Activation: cyanogen bromide Attachment: amino Spacer: 1 atom Ligand immobilized: 5-10 μmoles per ml Form: Suspension in 2.0 M NaCl containing 0.02% thimerosal	1 ml 13.30 5 ml 43.20 10 ml 71.45 25 ml 142.20
L-GLUTAMINE G 2884	Matrix: 4% beaded agarose Activation: cyanogen bromide Attachment: amino Spacer: 1 atom Ligand immobilized: 5-10 μmoles per ml Form: Suspension in 2.0 M NaCl	1 ml 13.30 5 ml 43.20 10 ml 71.45 25 ml 142.20
G 9644	Matrix: Cross-linked 4% beaded agarose Activation: epoxy Attachment: amino Spacer: 12 atoms Ligand immobilized: 0.5-1.0 μmole per ml Form: Suspension in 2.0 M NaCl containing 0.02% thimerosal	1 ml 13.30 5 ml 43.20 10 ml 71.45 25 ml 142.20
GLYCINE G 3009	Matrix: 4% beaded agarose Activation: cyanogen bromide Attachment: amino Spacer: 1 atom Ligand immobilized: 2-10 μmoles per ml Form: Suspension in 2.0 M NaCl containing 0.02% thimerosal	1 ml 14.50 5 ml 47.15 10 ml 78.05 25 ml 155.30
L-HISTIDINE H 3257	Matrix: 4% beaded agarose Activation: cyanogen bromide Attachment: amino Spacer: 1 atom Ligand immobilized: 2-10 μmoles per ml Form: Suspension in 2.0 M NaCl containing 0.02% thimerosal	1 ml 12.80 5 ml 41.35 10 ml 68.30 25 ml 135.75
H 0133	Matrix: Cross-linked 4% beaded agarose Activation: cyanogen bromide Attachment: amino Spacer: 1 atom Ligand immobilized: 10-20 μmoles per ml Form: Suspension in 0.5 M NaCl	1 ml 13.3 5 ml 43.2 10 ml 71.45 25 ml 142.2

PRODUCT NUMBER		
(Continuation of) **L-HISTIDINE** H 0767	Matrix: Cross-linked 4% beaded agarose Activation: epoxy Attachment: amino Spacer: 12 atoms Ligand immobilized: 1-2 μmoles Form: Suspension in 2.0 M NaCl thimerosal	
L-ISOLEUCINE I 4505	Matrix: 4% beaded agarose Activation: cyanogen bromide Attachment: amino Spacer: 1 atom Ligand immobilized: 3-7 μmoles Form: Suspension in 2.0 M NaCl thimerosal	
L-LEUCINE L 5506	Matrix: 4% beaded agarose Activation: cyanogen bromide Attachment: amino Spacer: 1 atom Ligand immobilized: 2-10 μmole Form: Suspension in 2.0 M NaCl	
L-LYSINE L 5631	Matrix: 4% beaded agarose Activation: cyanogen bromide Attachment: amino Spacer: 1 atom Ligand immobilized: 4-7 μmoles per ml Form: Suspension in 2.0 M NaCl thimerosal	
L 9268	Matrix: Cross-linked 4% beaded agarose Activation: epoxy Attachment: amino Spacer: 12 atoms Ligand immobilized: 1.5-2.5 μm Form: Suspension in 2.0 M NaCl thimerosal	
L 6132	Matrix: Sepharose 4B (pfs) Activation: cyanogen bromide Attachment: α-amino Spacer: 1 atom Ligand immobilized: approx. 4 μ Form: Lyophilized powder sta and dextran	
L-METHIONINE M 5010	Matrix: 4% beaded agarose Activation: cyanogen bromide Attachment: amino Spacer: 1 atom Ligand immobilized: 2-10 μmol Form: Suspension in 2.0 M NaCl	
L-PHENYLALANINE P 3018	Matrix: 4% beaded agarose Activation: cyanogen bromide Attachment: amino Spacer: 1 atom Ligand immobilized: 2-10 μmol Form: Suspension in 2.0 M NaCl	

For infor:

For shipping restrictions and related costs see page 5.

AMINO ACID RESINS (continued)

Left margin (partial, cut off):

s and enzymes is well...
otein separation.

		US $
	1 ml	15.60
arose	5 ml	49.70
bromide	10 ml	82.50
	25 ml	164.10

).5-1.0 µmole per ml
1 0.5 M NaCl, 0.01 M citrate,
.01% thimerosal

Inquire

	1 ml	13.30
agarose	5 ml	43.20
n bromide	10 ml	71.45
	25 ml	142.20

: 5-10 µmoles per ml
in 2.0 M NaCl containing 0.02%

agarose	1 ml	13.30
en bromide	5 ml	43.20
	10 ml	71.45
	25 ml	142.20

d: 5-10 µmoles

1 in 2.0 M NaCl

ed 4% beaded	1 ml	13.30
	5 ml	43.20
	10 ml	71.45
10	25 ml	142.20

ed: 0.5-1.0 µmole per ml
in in 2.0 M NaCl containing 0.02%

	1 ml	14.50
led agarose	5 ml	47.15
ogen bromide	10 ml	78.05
ino	25 ml	155.30

ized: 2-10 µmoles per ml
ion in 2.0 M NaCl containing 0.02%

aded agarose	1 ml	12.80
nogen bromide	5 ml	41.35
mino	10 ml	68.30
n	25 ml	135.75

ilized: 2-10 µmoles

sion in 2.0 M NaCl containing 0.02%

-linked 4% beaded	1 ml	13.30
	5 ml	43.20
yanogen bromide	10 ml	
amino	25 ml	

m
bilized: 10-20 µmoles per ml
nsion in 0.5 M NaCl

(Continued)

page 5.

Middle column

(Continuation of)
L-HISTIDINE

H 0767	Matrix: Cross-linked 4% beaded agarose	1 ml	12.80
		5 ml	41.00
	Activation: epoxy	10 ml	68.30
	Attachment: amino	25 ml	135.75

Spacer: 12 atoms
Ligand immobilized: 1-2 µmoles per ml
Form: Suspension in 2.0 M NaCl containing 0.02% thimerosal

I 4505	Matrix: 4% beaded agarose	1 ml	13.10
		5 ml	41.65
	Activation: cyanogen bromide	10 ml	68.65
	Attachment: amino	25 ml	136.25

L-ISOLEUCINE

Spacer: 1 atom
Ligand immobilized: 3-7 µmoles per ml
Form: Suspension in 2.0 M NaCl containing 0.02% thimerosal

L 5506	Matrix: 4% beaded agarose	1 ml	14.00
		5 ml	45.40
	Activation: cyanogen bromide	10 ml	75.25
	Attachment: amino	25 ml	149.55

L-LEUCINE

Spacer: 1 atom
Ligand immobilized: 2-10 µmoles per ml
Form: Suspension in 2.0 M NaCl

L-LYSINE

L 5631	Matrix: 4% beaded agarose	1 ml	15.70
	Activation: cyanogen bromide	5 ml	51.95
	Attachment: amino	10 ml	86.55
	Spacer: 1 atom	25 ml	173.00
	Ligand immobilized: 4-7 µmoles per ml	50 ml	311.60

Form: Suspension in 2.0 M NaCl containing 0.02% thimerosal

L 9268	Matrix: Cross-linked 4% beaded agarose	1 ml	15.70
		5 ml	51.95
	Activation: epoxy	10 ml	86.55
	Attachment: amino	25 ml	173.00

Spacer: 12 atoms
Ligand immobilized: 1.5-2.5 µmoles per ml
Form: Suspension in 2.0 M NaCl containing 0.02% thimerosal

C 6132	Matrix: Sepharose 4B (pfs)	1 g	46.45
	Activation: cyanogen bromide	5 g	154.75
	Attachment: α-amino	15 g	320.00

Spacer: 1 atom
Ligand immobilized: approx. 4 µmoles per ml
Form: Lyophilized powder stabilized with lactose and dextran

M 5010	Matrix: 4% beaded agarose	1 ml	13.30
		5 ml	43.20
	Activation: cyanogen bromide	10 ml	71.45
	Attachment: amino	25 ml	142.20

L-METHIONINE

Spacer: 1 atom
Ligand immobilized: 2-10 µmoles per ml
Form: Suspension in 2.0 M NaCl

C 3018	Matrix: 4% beaded agarose	1 ml	15.25
		5 ml	48.15
	Activation: cyanogen bromide	10 ml	92.35
	Attachment: amino	25 ml	184.00

L-PHENYLALANINE

Spacer: 1 atom
Ligand immobilized: 2-10 µmoles per ml
Form: Suspension in 2.0 M NaCl

Right column

P 3268	Matrix: 4% beaded agarose	1 ml	13.30
		5 ml	43.20
	Activation: cyanogen bromide	10 ml	71.45
	Attachment: amino	25 ml	142.20

L-PROLINE

Spacer: 1 atom
Ligand immobilized: 2-10 µmole per ml
Form: Suspension in 2.0 M NaCl containing 0.02% thimerosal

S 3881	Matrix: 4% beaded agarose	1 ml	13.30
		5 ml	43.20
	Activation: cyanogen bromide	10 ml	71.45
	Attachment: amino	25 ml	142.20

L-SERINE

Spacer: 1 atom
Ligand immobilized: 2-10 µmoles per ml
Form: Suspension in 2.0 M NaCl

T 0387	Matrix: 4% beaded agarose	1 ml	12.90
		5 ml	42.00
	Activation: cyanogen bromide	10 ml	69.40
	Attachment: amino	25 ml	138.15

L-THREONINE

Spacer: 1 atom
Ligand immobilized: 2-10 µmoles per ml
Form: Suspension in 2.0 M NaCl containing 0.02% thimerosal

T 3762	Matrix: 4% beaded agarose	1 ml	13.50
		5 ml	44.25
	Activation: cyanogen bromide	10 ml	72.40
	Attachment: amino	25 ml	144.15

D-TRYPTOPHAN

Spacer: 1 atom
Ligand immobilized: 1-3 µmoles per ml
Form: Suspension in 2.0 M NaCl

T 0137	Matrix: 4% beaded agarose	1 ml	13.50
		5 ml	44.25
	Activation: cyanogen bromide	10 ml	72.40
	Attachment: amino	25 ml	144.15

L-TRYPTOPHAN

Spacer: 1 atom
Ligand immobilized: 1-3 µmoles per ml
Form: Suspension in 2.0 M NaCl containing 0.02% thimerosal

L-TYROSINE

T 0262	Matrix: 4% beaded agarose	1 ml	14.85
	Activation: cyanogen bromide	5 ml	45.00
	Attachment: amino	10 ml	74.50
	Spacer: 1 atom	25 ml	148.50

Ligand immobilized: 5-10 µmoles per ml
Form: Suspension in 2.0 M NaCl containing 0.02% thimerosal

T 8016	Matrix: 4% beaded agarose	1 ml	17.40
	Activation: epoxy	5 ml	55.90
	Attachment: amino	25 ml	185.20

Spacer: 12 atoms
Ligand immobilized: approx. 20 µmoles per ml
Form: Suspension in 0.15 M sodium phosphate, pH 6.8, containing 0.02% sodium azide

V 5877	Matrix: 4% beaded agarose	1 ml	13.10
		5 ml	41.65
	Activation: cyanogen bromide	10 ml	68.65
	Attachment: amino	25 ml	136.25

L-VALINE

Spacer: 1 atom
Ligand immobilized: 3-7 µmoles per ml
Form: Suspension in 2.0 M NaCl containing 0.02% thimerosal

BIOCHEMICALS
ORGANIC COMPOUNDS

DIAGNOSTIC REAGENTS

applicant

TOLL FREE ORDERING

ORDERING HOURS
Monday through Friday—24 Hours a day
Saturday and Holidays—6:00 a.m. to 5:00 p.m. (CST)
Sunday—6:00 a.m. to 1:00 p.m. (CST)

TELEPHONE USA / Canada **800-325-3010**

Outside USA / Canada Call Collect **314-771-5750**

FAX USA / Canada **800-325-5052**

Outside USA / Canada **314-771-5757**
For TWX, TELEX, Cable or Collect FAX calls see page 3.

FOR BULK INQUIRIES CALL SAF
See inside back cover for details.

Customer / Technical Service

Customer Service/Technical Service telephone numbers located
on card inside front cover and on page 3.

Should you have any problem getting excellent service at the above numbers, please call
COLLECT (anytime day or night, including weekends and holidays) 314-531-9500.

SIGMA®

ALPHABETICAL LIST

BIOACTIVE PEPTIDES

IMMUNO- CHEMICALS

MOLECULAR BIOLOGY

TISSUE CULTURE

OTHER PRODUCT GROUPS

EQUIPMENT BOOKS AND SUPPLIES

DIAGNOSTIC KITS AND REAGENTS

BIOCHEMICALS
ORGANIC COMPOUNDS

DIAGNOSTIC REAGENTS

file copy applicant

TOLL FREE ORDERING

ORDERING HOURS
Monday through Friday—24 Hours a day
Saturday and Holidays—6:00 a.m. to 5:00 p.m. (CST)
Sunday—6:00 a.m. to 1:00 p.m. (CST)

TELEPHONE USA / Canada **800-325-3010**

Outside USA / Canada Call Collect **314-771-5750**

FAX USA / Canada **800-325-5052**

Outside USA / Canada **314-771-5757**
For TWX, TELEX, Cable or Collect FAX calls see page 3.

FOR BULK INQUIRIES CALL SAF
See inside back cover for details.

Customer/Technical Service

Customer Service/Technical Service telephone numbers located
on card inside front cover and on page 3.

Should you have any problem getting excellent service at the above numbers, please call
COLLECT (anytime day or night, including weekends and holidays) 314-531-9500.

SIGMA®

ALPHABETICAL LIST

BIOACTIVE PEPTIDES

IMMUNO-CHEMICALS

MOLECULAR BIOLOGY

TISSUE CULTURE

OTHER PRODUCT GROUPS

EQUIPMENT BOOKS AND SUPPLIES

DIAGNOSTIC KITS AND REAGENTS

AFFINITY CHROMATOGRAPHY APPLICATIONS INDEX

The table is intended as a general guide to selecting an affinity chromatography medium. Where a medium is suggested for a very specific application, literature references are given; those references should be consulted for appropriate chromatography conditions and limitations of the method.

References are also given to support more general guidelines, such as ovomucoid-agarose for lectins. Where specific references to your particular purification problem do not exist, the general suggestions may help you select an appropriate medium. Such applications must be viewed as investigatory, however; you should expect to have to do some or many trials to arrive at the best combination of medium and conditions.

For Purification/Binding of	Suggested ligand
adenosine monophosphate	acriflavine[2]
albumin	cholic acid[3]
amino acids, D and L	albumin[**]
antibodies	protein A; protein G; anti-Ig; serum proteins
avidin-labeled compounds	biotin
bilirubin	albumin[47]
biotin and biotin-labeled compounds	avidin; streptavidin
cells	lectins on macroagarose (agarose macrobeads)
halophilic bacteria	ω-aminohexyl[118]
HeLa cells	lens culinaris lectin[64]
human erythrocytes	anti-rabbit IgG[**]
mouse thymocytes	protein A[77]
chorionic gonadotropin	concanavalin A[6]
coagulation factors	heparin[13]
coagulation factor II	sulfated dextran[14]; heparin[13]
plasminogen	L-lysine[62]
plasminogen activators	p-aminobenzamidine[83]
thrombin	m-aminobenzamidine[**]; p-aminobenzamidine[43]
dihydrofolate reductase	Cibacron blue 3GA[116]
endotoxins (pyrogens)	histamine[15]; polymyxin B[16]
enzymes	amino acids; dyes; nucleotides/cofactors
N-acetylhexosaminidase	concanavalin A[6]
adenosine deaminase	ω-aminohexyl[122]
adenosine kinase	adenosine 5'-monophosphate[75]
alcohol dehydrogenase	Cibacron blue 3GA[116]; adenosine 5'-monophosphate[78]
alkaline phosphatase	L-histidyldiazobenzylphosphonic acid[20]
aminoacyl tRNA-synthetase	ω-aminohexyl[119]
aminopeptidase	ω-aminohexyl[123]
anhydrochymotrypsin	trypsin inhibitor[26]
anthranilate synthetase	L-tryptophan[88]
carbohydrate-metabolizing enzymes	carbohydrates
carbonic anhydrase	p-aminomethylbenzenesulfonamide[22]
carboxypeptidase A	D-tryptophan[84]
casein kinase	α-casein[82]; heparin[83]
cholesterol oxidase	cholesteryl hemisuccinate[23]
α-chymotrypsin	4-phenylbutylamine[24]
cytochrome oxidase	cytochrome c[25]
DAHP synthetase, Tyr-sensitive	L-tyrosine[66]
dehydrogenases	nucleotides; NAD; NADP; Cibacron blue 3GA[116]
	reactive red 120[110]
detoxifying enzymes	glutathione
dihydrofolate reductase	methotrexate[30]; Cibacron blue 3GA[116]
dihydroneopterin triphosphate synthetase	guanosine 5'-triphosphate[74]
enterokinase	p-aminobenzamidine[43]
α-galactosidase	α(+)melibiose[106]
β-galactosidase	p-aminobenzyl 1-thio-β-D-galactopyranoside[31]
	p-aminophenyl β-D-thiogalactopyranoside[31]
	p-aminobenzyl 1-thio-β-D-galactopyranoside[31]
galactosyltransferase	α-lactalbumin[33]; N-acetyl-D-glucosamine[88]
glucuronidase	saccharolactone[108]
glutamate dehydrogenase	ω-aminopentyl
glutamate synthetase	Cibacron blue 3GA[6]
glutathione S-transferase	cholic acid[34]
glyceraldehyde 3-phosphate dehydrogenase	reactive blue 72; pentyl
glyceraldehyde dehydrogenase	adenosine 5'-monophosphate[77]
glycogen phosphorylase b	ω-aminohexyl[121]; butyl[114]
glycogen synthetase	ω-aminobutyl[131]; uridine 5'-diphosphate[77]
hexokinase	N-acetyl-D-glucosamine[108]
histone kinase	heparin[163]
kallikrein	trypsin inhibitor[31]
kinases	nucleotides; ATP; Cibacron blue 3GA[18]
lactic dehydrogenase	β-nicotinamide adenine dinucleotide[73]

Left margin fragments:

ng the buf-
lumn at a
used to elute
ng effects
required for
order to have
capacity

US $

25 ml   43.45
.00 ml 120.60
:00 ml 201.00

'fer 96

ive com-
most listings

we offer a com-

Press,

AMINO ACID RESINS

Immobilized amino acids are very versatile affinity media. Their use in the isolation of proteins and enzymes is well established. Recently, there has been increasing interest in the use of these resins for serum protein separation.

General References:
Deutsch, D.G. and Mertz, E.T., Proc. Fed. Amer. Soc. Exp. Biol., **29**, 647 (1970).
Deutsch, D.G. and Mertz, E.T., Science, **170**, 1095 (1970).
Vverto, M. and Vaheki, A., Biochem. J., **183**, 331 (1979).

PRODUCT NUMBER		US $
D-ALANINE	1 ml	13.30
A 3435 Matrix: Cross-linked 4% beaded agarose	5 ml	43.20
	10 ml	71.45
Activation: cyanogen bromide		
Attachment: amino		
Spacer: 1 atom		
Ligand immobilized: 2-8 µmoles per ml		
Form: Suspension in 2.0 M NaCl containing 0.02% thimerosal		
L-ALANINE	1 ml	13.30
A 3519 Matrix: 4% beaded agarose	5 ml	43.20
Activation: cyanogen bromide	10 ml	71.45
Attachment: amino	25 ml	142.20
Spacer: 1 atom		
Ligand immobilized: 2-7 µmoles per ml		
Form: Suspension in 2.0 M NaCl containing 0.02% thimerosal		
L-ARGININE		
Matrix: 4% beaded agarose	1 ml	13.90
A 1018 Activation: cyanogen bromide	5 ml	44.70
Attachment: amino	10 ml	74.25
Spacer: 1 atom	25 ml	148.20
Ligand immobilized: 5-10 µmoles per ml		
Form: Suspension in 2.0 M NaCl containing 0.02% thimerosal		
Matrix: Cross-linked 4% beaded	1 ml	10.80
A 8405 agarose	5 ml	34.95
Activation: epoxy	10 ml	57.80
Attachment: amino		
Spacer: 12 atoms		
Ligand immobilized: 1-3 µmoles per ml		
Form: Lyophilized powder stabilized with lactose		
Swelling: 1 g swells to approx. 12 ml		
L-ASPARAGINE		
Matrix: 4% beaded agarose	1 ml	13.30
A 1143 Activation: cyanogen bromide	5 ml	43.20
Attachment: amino	10 ml	71.45
Spacer: 1 atom	25 ml	142.20
Ligand immobilized: 5-10 µmoles per ml		
Form: Suspension in 2.0 M NaCl containing 0.02% thimerosal		
Matrix: Cross-linked 4% beaded	1 ml	13.30
A 2047 agarose	5 ml	43.20
Activation: epoxy	10 ml	71.45
Attachment: amino	25 ml	142.20
Spacer: 12 atoms		
Ligand immobilized: 1-2 µmoles per ml		
Form: Suspension in 2.0 M NaCl containing 0.02% thimerosal		
L-ASPARTIC ACID	1 ml	21.20
A 3394 Matrix: 4% beaded agarose	5 ml	68.50
Activation: cyanogen bromide	10 ml	113.80
Attachment: amino	25 ml	227.15
Spacer: 1 atom		
Ligand immobilized: 5-10 µmoles per ml		
Form: Suspension in 2.0 M NaCl containing 0.02% thimerosal		

PRODUCT NUMBER		US $
L-CYSTEINE	1 ml	15.60
C 7896 Matrix: 4% beaded agarose	5 ml	49.70
Activation: cyanogen bromide	10 ml	82.50
Attachment: amino	25 ml	164.10
Spacer: 1 atom		
Ligand immobilized: 0.5-1.0 µmole per ml		
Form: Suspension in 0.5 M NaCl, 0.01 M citrate pH 4.5, containing 0.01% thimerosal		
L-CYSTINE		Inquire
C 9266 Matrix: agarose		
L-GLUTAMIC ACID	1 ml	13.30
G 2759 Matrix: 4% beaded agarose	5 ml	43.20
Activation: cyanogen bromide	10 ml	71.45
Attachment: amino	25 ml	142.20
Spacer: 1 atom		
Ligand immobilized: 5-10 µmoles per ml		
Form: Suspension in 2.0 M NaCl containing 0.02% thimerosal		
L-GLUTAMINE		
Matrix: 4% beaded agarose	1 ml	13.30
G 2884 Activation: cyanogen bromide	5 ml	43.20
Attachment: amino	10 ml	71.45
Spacer: 1 atom	25 ml	142.20
Ligand immobilized: 5-10 µmoles per ml		
Form: Suspension in 2.0 M NaCl		
Matrix: Cross-linked 4% beaded	1 ml	13.30
G 9644 agarose	5 ml	43.20
Activation: epoxy	10 ml	71.45
Attachment: amino	25 ml	142.20
Spacer: 12 atoms		
Ligand immobilized: 0.5-1.0 µmole per ml		
Form: Suspension in 2.0 M NaCl containing 0.02% thimerosal		
GLYCINE	1 ml	14.50
G 3009 Matrix: 4% beaded agarose	5 ml	47.15
Activation: cyanogen bromide	10 ml	78.05
Attachment: amino	25 ml	155.35
Spacer: 1 atom		
Ligand immobilized: 2-10 µmoles per ml		
Form: Suspension in 2.0 M NaCl containing 0.02% thimerosal		
L-HISTIDINE		
Matrix: 4% beaded agarose	1 ml	12.80
H 3257 Activation: cyanogen bromide	5 ml	41.35
Attachment: amino	10 ml	68.30
Spacer: 1 atom	25 ml	135.85
Ligand immobilized: 2-10 µmoles per ml		
Form: Suspension in 2.0 M NaCl containing 0.02% thimerosal		
Matrix: Cross-linked 4% beaded	1 ml	13.30
H 0133 agarose	5 ml	43.20
Activation: cyanogen bromide	10 ml	71.45
Attachment: amino	25 ml	142.20
Spacer: 1 atom		
Ligand immobilized: 10-20 µmoles per ml		
Form: Suspension in 0.5 M NaCl		

(Continued)

PRODUCT NUMBER		US $
(Continuation of) **L-HISTIDINE**		
Matrix: Cross-linked 4% beaded		
H 0767 agarose		
Activation: epoxy		
Attachment: amino		
Spacer: 12 atoms		
Ligand immobilized: 1-2 µmol		
Form: Suspension in 2.0 M N thimerosal		
L-ISOLEUCINE		
I 4505 Matrix: 4% beaded agarose		
Activation: cyanogen bromide		
Attachment: amino		
Spacer: 1 atom		
Ligand immobilized: 3-7 µmol		
Form: Suspension in 2.0 M thimerosal		
L-LEUCINE		
L 5506 Matrix: 4% beaded agarose		
Activation: cyanogen bromide		
Attachment: amino		
Spacer: 1 atom		
Ligand immobilized: 2-10 µm		
Form: Suspension in 2.0 M.N		
L-LYSINE		
Matrix: 4% beaded agarose		
L 5631 Activation: cyanogen bromide		
Attachment: amino		
Spacer: 1 atom		
Ligand immobilized: 4-7 µmo per ml		
Form: Suspension in 2.0 M thimerosal		
Matrix: Cross-linked 4% bea		
L 9268 agarose		
Activation: epoxy		
Attachment: amino		
Spacer: 12 atoms		
Ligand immobilized: 1.5-2.5		
Form: Suspension in 2.0 M thimerosal		
Matrix: Sepharose 4B (pfs)		
L 6132 Activation: cyanogen bromid		
Attachment: α-amino		
Spacer: 1 atom		
Ligand immobilized: approx.		
Form: Lyophilized powder and dextran		
L-METHIONINE		
M 5010 Matrix: 4% beaded agarose		
Activation: cyanogen bromic		
Attachment: amino		
Spacer: 1 atom		
Ligand immobilized: 2-10 µr		
Form: Suspension in 2.0 M		
L-PHENYLALANINE		
P 3018 Matrix: 4% beaded agarose		
Activation: cyanogen bromi		
Attachment: amino		
Spacer: 1 atom		
Ligand immobilized: 2-10 µl		
Form: Suspension in 2.0 M		

For inf

For shipping restrictions and related costs see page 5.

AMINO ACID RESINS (continued)

Left-margin fragments (from facing page):

s and enzymes is well
otein separation.

	US $
	1 ml 15.60
arose	5 ml 49.70
bromide	10 ml 82.50
	25 ml 164.10

.5-1.0 µmole per ml
0.5 M NaCl, 0.01 M citrate
01% thimerosal

	Inquire
	1 ml 13.30
agarose	5 ml 43.20
bromide	10 ml 71.45
	25 ml 142.20

5-10 µmoles per ml
in 2.0 M NaCl containing 0.02%

	1 ml 13.30
agarose	5 ml 43.20
en bromide	10 ml 71.45
	25 ml 142.20

d: 5-10 µmoles
in 2.0 M NaCl
ed 4% beaded · 1 ml 13.30, 5 ml 43.20, 10 ml 71.45, 25 ml 142.20

ed: 0.5-1.0 µmole per ml
n in 2.0 M NaCl containing 0.02%

	1 ml 14.50
led agarose	5 ml 47.15
ogen bromide	10 ml 78.05
ino	25 ml 155.30

ized: 2-10 µmoles per ml
ion in 2.0 M NaCl containing 0.0

	1 ml 12
aded agarose	5 ml 41
inogen bromide	10 ml 68
mino	25 ml 135
n	

ilized: 2-10 µmoles
ision in 2.0 M NaCl containing 0.

-linked 4% beaded · 1 ml, 5 ml, 10 ml, 25 ml
yanogen bromide
amino
om
ibilized: 10-20 µmoles per ml
ension in 0.5 M NaCl
(Continue

page 5.

PRODUCT NUMBER / US $

(Continuation of) L-HISTIDINE
H 0767 — Matrix: Cross-linked 4% beaded agarose — 1 ml 12.80, 5 ml 41.00, 10 ml 68.30, 25 ml 135.75
Activation: epoxy
Attachment: amino
Spacer: 12 atoms
Ligand immobilized: 1-2 µmoles per ml
Form: Suspension in 2.0 M NaCl containing 0.02% thimerosal

L-ISOLEUCINE
I 4505 — Matrix: 4% beaded agarose — 1 ml 13.10, 5 ml 41.65, 10 ml 68.65, 25 ml 136.25
Activation: cyanogen bromide
Attachment: amino
Spacer: 1 atom
Ligand immobilized: 3-7 µmoles per ml
Form: Suspension in 2.0 M NaCl containing 0.02% thimerosal

L-LEUCINE
L 5506 — Matrix: 4% beaded agarose — 1 ml 14.00, 5 ml 45.40, 10 ml 75.25, 25 ml 149.55
Activation: cyanogen bromide
Attachment: amino
Spacer: 1 atom
Ligand immobilized: 2-10 µmoles per ml
Form: Suspension in 2.0 M NaCl

L-LYSINE
L 5631 — Matrix: 4% beaded agarose — 1 ml 15.70, 5 ml 51.95, 10 ml 86.55, 25 ml 173.00, 50 ml 311.60
Activation: cyanogen bromide
Attachment: amino
Spacer: 1 atom
Ligand immobilized: 4-7 µmoles per ml
Form: Suspension in 2.0 M NaCl containing 0.02% thimerosal

L 9268 — Matrix: Cross-linked 4% beaded agarose — 1 ml 15.70, 5 ml 51.95, 10 ml 86.55, 25 ml 173.00
Activation: epoxy
Attachment: amino
Spacer: 12 atoms
Ligand immobilized: 1.5-2.5 µmoles per ml
Form: Suspension in 2.0 M NaCl containing 0.02% thimerosal

C6132 — Matrix: Sepharose 4B (pfs) — 1 g 46.45, 5 g 154.75, 15 g 320.00
Activation: cyanogen bromide
Attachment: α-amino
Spacer: 1 atom
Ligand immobilized: approx. 4 µmoles per ml
Form: Lyophilized powder stabilized with lactose and dextran

L-METHIONINE
M 5010 — Matrix: 4% beaded agarose — 1 ml 13.30, 5 ml 43.20, 10 ml 71.45, 25 ml 142.20
Activation: cyanogen bromide
Attachment: amino
Spacer: 1 atom
Ligand immobilized: 2-10 µmoles per ml
Form: Suspension in 2.0 M NaCl

L-PHENYLALANINE
P 3018 — Matrix: 4% beaded agarose — 1 ml 15.25, 5 ml 48.15, 10 ml 92.35, 25 ml 184.00
Activation: cyanogen bromide
Attachment: amino
Spacer: 1 atom
Ligand immobilized: 2-10 µmoles per ml
Form: Suspension in 2.0 M NaCl

L-PROLINE
P 3268 — Matrix: 4% beaded agarose — 1 ml 13.30, 5 ml 43.20, 10 ml 71.45, 25 ml 142.20
Activation: cyanogen bromide
Attachment: amino
Spacer: 1 atom
Ligand immobilized: 2-10 µmole per ml
Form: Suspension in 2.0 M NaCl containing 0.02% thimerosal

L-SERINE
S 3881 — Matrix: 4% beaded agarose — 1 ml 13.30, 5 ml 43.20, 10 ml 71.45, 25 ml 142.20
Activation: cyanogen bromide
Attachment: amino
Spacer: 1 atom
Ligand immobilized: 2-10 µmoles per ml
Form: Suspension in 2.0 M NaCl

L-THREONINE
T 0387 — Matrix: 4% beaded agarose — 1 ml 12.90, 5 ml 42.00, 10 ml 69.40, 25 ml 138.15
Activation: cyanogen bromide
Attachment: amino
Spacer: 1 atom
Ligand immobilized: 2-10 µmoles per ml
Form: Suspension in 2.0 M NaCl containing 0.02% thimerosal

D-TRYPTOPHAN
T 3762 — Matrix: 4% beaded agarose — 1 ml 13.50, 5 ml 44.25, 10 ml 72.40, 25 ml 144.15
Activation: cyanogen bromide
Attachment: amino
Spacer: 1 atom
Ligand immobilized: 1-3 µmoles per ml
Form: Suspension in 2.0 M NaCl

L-TRYPTOPHAN
T 0137 — Matrix: 4% beaded agarose — 1 ml 13.50, 5 ml 44.25, 10 ml 72.40, 25 ml 144.15
Activation: cyanogen bromide
Attachment: amino
Spacer: 1 atom
Ligand immobilized: 1-3 µmoles per ml
Form: Suspension in 2.0 M NaCl containing 0.02% thimerosal

L-TYROSINE
T 0262 — Matrix: 4% beaded agarose — 1 ml 14.85, 5 ml 45.00, 10 ml 74.50, 25 ml 148.50
Activation: cyanogen bromide
Attachment: amino
Spacer: 1 atom
Ligand immobilized: 5-10 µmoles per ml
Form: Suspension in 2.0 M NaCl containing 0.02% thimerosal

T 8016 — Matrix: 4% beaded agarose — 1 ml 17.40, 5 ml 55.90, 25 ml 185.20
Activation: epoxy
Attachment: amino
Spacer: 12 atoms
Ligand immobilized: approx. 20 µmoles per ml
Form: Suspension in 0.15 M sodium phosphate, pH 6.8, containing 0.02% sodium azide

L-VALINE
V 5877 — Matrix: 4% beaded agarose — 1 ml 13.10, 5 ml 41.65, 10 ml 68.65, 25 ml 136.25
Activation: cyanogen bromide
Attachment: amino
Spacer: 1 atom
Ligand immobilized: 3-7 µmoles per ml
Form: Suspension in 2.0 M NaCl containing 0.02% thimerosal

Please provide the following references

L1   ANSWER 13 OF 17   MEDLINE
AN   97016817   MEDLINE
DN   97016817   PubMed ID: 8863443
TI   Organization and phylogenetic interrelationships of genes encoding
     components of the botulinum toxin complex in proteolytic Clostridium
     botulinum types A, B, and F: evidence of chimeric sequences in the gene
     encoding the nontoxic nonhemagglutinin component.
AU   East A K; Bhandari M; Stacey J M; Campbell K D; Collins M D
CS   Department of Microbiology, Institute of Food Research, Reading,
     Berkshire, United Kingdom.. alison.east@bbsrc.ac.uk
SO   INTERNATIONAL JOURNAL OF SYSTEMATIC BACTERIOLOGY, (1996 Oct) 46 (4)
     1105-12.
     Journal code: 0042143. ISSN: 0020-7713.
CY   United States
DT   Journal; Article; (JOURNAL ARTICLE)
LA   English
FS   Priority Journals
OS   GENBANK-X78230; GENBANK-X87974; GENBANK-X92973; GENBANK-X96491;
     GENBANK-X96492; GENBANK-X96493; GENBANK-X96494
EM   199611
ED   Entered STN: 19961219
     Last Updated on STN: 19980206
     Entered Medline: 19961115

L1   ANSWER 14 OF 17   MEDLINE
AN   94297488   MEDLINE
DN   94297488   PubMed ID: 7764998
TI   Conserved structure of genes encoding components of botulinum
     neurotoxin complex M and the sequence of the gene coding for the
     nontoxic component in nonproteolytic Clostridium botulinum type
     F.
AU   East A K; Collins M D
CS   Department of Microbiology, Institute of Food Research, Reading
     Laboratory, UK.
SO   CURRENT MICROBIOLOGY, (1994 Aug) 29 (2) 69-77.
     Journal code: 7808448. ISSN: 0343-8651.
CY   United States
DT   Journal; Article; (JOURNAL ARTICLE)
LA   English
FS   Biotechnology
EM   199408
ED   Entered STN: 19950809
     Last Updated on STN: 19970203
     Entered Medline: 19940810

No match

L1   ANSWER 15 OF 17   MEDLINE
AN   93012902   MEDLINE
DN   93012902   PubMed ID: 1398040
TI   Sequence of the gene encoding type F neurotoxin of
     Clostridium botulinum.
AU   East A K; Richardson P T; Allaway D; Collins M D; Roberts T A; Thompson D
     E

1

Conserved Structure of Genes Encoding Components of Botulinum Neurotoxin Complex M and the Sequence of the Gene Coding for the Nontoxic Component in Nonproteolytic *Clostridium botulinum* Type F

Alison K. East, Matthew D. Collins

Department of Microbiology, Institute of Food Research, Reading Laboratory, Earley Gate, Whiteknights Road, Reading RG6 2EF, UK

Abstract. For investigation of the genes of proteins associated in vivo with botulinum neurotoxin (BoNT), polymerase chain reaction (PCR) experiments were carried out with oligonucleotide primers designed to regions of the nontoxic-nonhemagglutinin (*NTNH*) gene of *Clostridium botulinum* type C. The primers were used to amplify a DNA fragment from genomic DNA of *C. botulinum* types A, B, E, F, G and toxigenic strains of *Clostridium barati* and *Clostridium butyricum*. The amplified product from all of these strains hybridized with an internal oligonucleotide probe, whereas all nontoxigenic clostridia tested gave no PCR product and showed no reaction with the probe. The *NTNH* gene was shown to be located upstream of the gene encoding BoNT, thereby revealing a conserved structure for genes encoding the proteins of the M complex of the progenitor botulinum toxin in these organisms. The sequence of the *NTNH* gene of nonproteolytic *C. botulinum* type F was determined by PCR amplification and sequencing of overlapping cloned fragments. NTNH/F showed 71% and 61% identity with NTNH of *C. botulinum* type E and type C respectively.

Botulinum neurotoxin is the causative agent of botulism. It is produced by strains of four different physiological groups (I–IV), all designated *Clostridium botulinum* [7], and some strains of *C. barati* and *C. butyricum* [9, 10]. Strains of *C. botulinum* are classified into seven types, A to G, depending on the antigenicity of BoNT produced [7]. The toxin is produced by these organisms as a progenitor toxin complex, which is found in three forms: (i) M (rmm ~ 300,000) consisting of BoNT (rmm ~ 150,000) and a nontoxic protein component of approximately equal size, (ii) L (rmm ~ 500,000), and (iii) LL (rmm ~ 900,000). The larger L and LL complexes have hemagglutinin activity and have been purified from group I (proteolytic) *C. botulinum* producing BoNT/A and BoNT/B [14] and group III *C. botulinum* producing BoNT/C (L complex only) [17]. In contrast, group II (nonproteolytic) strains of *C. botulinum* that produce BoNT/E have no hemagglutinin activity, and only the M form

of the complex has been detected [14]. The situation is similar for type A strains, which cause infant botulism, with only the M form being observed, with no hemagglutinin activity or larger complexes present [12].

The gene encoding the NTNH component, which together with BoNT forms the M complex, has been sequenced for *C. botulinum* type C [17]. More recently, the same workers published the sequence of *ent*-120 from *C. botulinum* type E [5] and the same gene of toxigenic *C. butyricum* [6], both of which show high homology with *NTNH/C*. We have reported the DNA sequence encoding the C-terminus of a protein showing homology with NTNH/C, upstream of *BoNT* gene in *C. botulinum* types A, E, and F (group II), *C. butyricum* type E, and *C. barati* type F [16]. In this study we show that the gene encoding this component of the progenitor toxin complex is present in all BoNT-producing strains, irrespective of immunological toxin type, and is absent from strains not producing BoNT. Furthermore, the *NTNH* gene is present in

Correspondence to: A.K. East

Table 1. Details of species and strains tested

Species	Strain[a]	Toxin	Physiological group
C. botulinum	NCTC 7272	A	I
	Kyoto-F	A(inf)	I
	NCTC 7273	B	I
	Eklund 17B (ATCC 25765)	B	II
	Hobbs FT50 (Colworth 194)	B	II
	Eklund 2B (Colworth 229)	B	II
	Colworth 151	B	II
	Scott 2129B	B	II
	Beluga (Colworth)	E	II
	Hazen 36208 (ATCC 9564)	E	II
	Sebald P34 (Colworth 230)	E	II
	VH (Dolman)	E	II
	Langeland (NCTC 10281)	F	I
	Eklund 202F (ATCC 23387)	F	II
	Hobbs FT10 (Colworth 187)	F	II
	Craig 610 (NCIMB 4305)	F	II
	ATCC 27322	G	IV
C. botulinum (nontoxigenic)	VPI 4088 (Nontoxic type E, NTE)	—	II
	VPI 10428 (NTE)	—	II
	VPI 2093-1 (NTE)	—	II
	VPI 3158 (NTE)	—	II
	4672 U-1 (NTB)	—	II
C. butyricum	ATCC 43755	E	—
C. barati	ATCC 43756	F	—
C. hastiforme	DSM 5675	—	—

[a] ATCC, American Type Culture Collection; DSM, Deutsche Sammlung von Mikroorganismen; NCIMB, National Collection of Industrial and Marine Bacteria; NCTC, National Collection of Type Cultures; VPI, Virginia Polytechnic Institute and State University.

a common location upstream of the *BoNT* gene. We also present the sequence of the gene encoding the NTNH component from group II *C. botulinum* type F.

Materials and Methods

Bacterial strains. The strains of clostridia used in this study are listed in Table 1.

Isolation of DNA. DNA was isolated from clostridial strains as described previously [4]. Plasmid DNA was purified from *E. coli* with a Magic Mini-Prep kit (Promega, Southampton, UK) or by a modification of the alkaline lysis method [13].

PCR amplification. Oligonucleotides were synthesised on an Applied Biosystems DNA synthesiser (MODEL 391, Warrington, UK), and PCR was performed with a Biometra thermal cycler (Maidstone, UK). PCR was performed as described previously [4] with 25 cycles of: 92°C for 1 min; 37°C for 1 min; 58°C for 5 min, and the product held at 4°C. Template DNA was at a final concentration of 1 ng/µl, except for inverse PCR, where 4 ng/µl was used. All primers were used at a final concentration of 4 ng/µl. The position and sequence of the primers used are shown in Figs. 1 and 3.

Cloning and transformation. Cloning of PCR products was as described previously [4] or by use of a TA cloning kit (Invitrogen, Witney, UK) according to the manufacturer's instructions.

Sequence determination. DNA sequencing was carried out as described previously [4]. Analysis of sequence data was carried out with Wisconsin Molecular Biology software [3] on a VAX computer. The sequence was determined on two cloned fragments derived from different PCR experiments. Where two clones differed in sequence, as they did in three positions, a third, independently amplified fragment was cloned and sequenced.

Southern blot analysis. Amplified DNA fragments (~1 µg) separated on a 0.8% agarose gel were transferred to Hybond N+ nylon membrane (Amersham International, Amersham, UK) by a modification of the method of Southern [13]. The DNA was fixed by placing the membrane on 3MM Whatman paper soaked in 0.4 M NaOH for 20 min. Oligonucleotide probe NN3 (5'-TTTAGTTTCT-TAGATCAATGGTGG-3') was labeled and detected with an Enhanced Chemiluminescence (ECL) kit (Amersham International), following the manufacturer's instructions. Blots were hybridized at 42°C for 2 h and washed at 42°C or 48°C in 1 × SSC, 0.1% (wt/vol) SDS for 30 min. After the addition of detection agents, the blot was exposed to X-ray film for between 30 s and 2 min.

Results and Discussion

Examination of clostridial strains for presence of *NTNH* gene. The presence of sequences similar to the *NTNH* gene was detected with primers NN1 and NN2. NN1 was based on the sequence of *NTNH/C* gene and N-terminal protein sequence data [14], and NN2 was designed with the sequences of *NTNH/C* gene [18] and the region upstream of the gene encoding *BoNT/F*. PCR was performed with primers NN1 and NN2 and template DNA prepared from the clostridial strains shown in Table 1 (Fig. 1). Only strains that produce BoNT gave a DNA fragment (approx. 1.9 kb in size) in the reaction (Fig. 2). Gels of these fragments were blotted and probed with oligonucleotide NN3, designed to the sequence FS-FLDQWW (amino acids 666 to 673) of NTNH/C [18]. This region was considered likely to be conserved, and the two tryptophan residues minimized the need for degenerate oligonucleotides. All of the amplified bands hybridized to probe NN3, but to differing extents, presumably reflecting the degree of identity between the probe and the target DNA sequences (Fig. 2).

Two strains producing BoNT/A were tested by PCR with NN1 and NN2: NCTC 7272 and Kyoto. The latter, isolated from a case of infant botulism, produces BoNT with a heavy chain of altered mobility on SDS gel electrophoresis [12]. A very faint PCR product was obtained with primers NN1 and NN2 for strain NCTC 7272, which hybridized with NN3 (Fig. 2), but none was visible with Kyoto DNA as a template, either as a PCR product after agarose
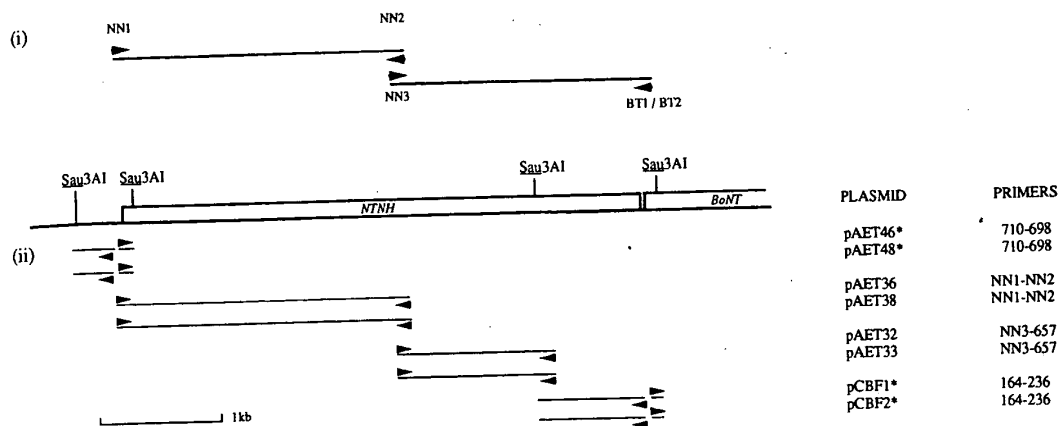
Fig. 1. Diagrammatic representation of *NTNH* gene showing (i) position of primers used in PCR and probing experiments and (ii) strategy for cloning fragments amplified by PCR to determine the sequence of *NTNH/F*. The plasmids marked * were obtained by cloning fragments obtained by inverse PCR, digesting the chromosomal DNA with *Sau*3AI in each case before ligation and use as the template in the reaction. The sequence of primers used in PCR/probing experiments was as follows; NN1: 5'-ATA(A/G)ATTC(A/T)CCAGTGGA(T/C)AA-3', NN2: 5'-ATTTTCTTGTGCTAATATTGA-3', NN3: 5'-TTAGTTTCTTAGATCAATGGTGG-3', BT1: 5'-AT(A/T)TCT(T/C)TC(A/T)-GGAATTA(T/C)CCAAATA-3', BT2: 5'-TCT(T/C)TC(A/T)GGAATTA(T/C)CCAAATATT-3'. The position of binding of all primers is indicated in Fig. 4.

electrophoresis or by probing with NN3 (Fig. 2). However, with primers designed to a different region of the gene (NN3 and BT1, Fig. 1 (i)), both type A strains gave a PCR product visible after gel electrophoresis (Fig. 3). Sequencing PCR products amplified from the region upstream of *BoNT/A* (Kyoto) has revealed an orf showing homology with *NTNH/C* [A. Willems, A.K. East, and M.D. Collins, unpublished results]. In addition to showing different properties in gel electrophoresis, the two type A toxins also differed in the antigenicity of the second component of the M-complex, i.e., NTNH [15]. The observed differences in immunogenicity of the two NTNH/A proteins are probably owing to variation in the sequence of their genes. Nucleotide changes at one or both of the primer (NN1 and/or NN2) binding sites in strain Kyoto may account for the PCR results.

Strains of proteolytic and nonproteolytic *C. botulinum* producing BoNT/B gave a PCR fragment that hybridized strongly with probe NN3 (Fig. 2). By contrast, a strain 4672U-1, phenotypically resembling nonproteolytic *C. botulinum* type B but producing no toxin, gave no PCR product [2]. Within a single BoNT immunological group, the efficiency of hybridization of probe NN3 seemed to be approximately the same (Fig. 2), possibly reflecting high sequence similarity in the *NTNH* genes of strains of a particular toxin type. In the case of *BoNT* genes, sequence determination of an ~ 1-kb region encoding part of the heavy chain, from several strains of the same immunological toxin type, showed <0.2% sequence variation [2].

DNA amplified from strains of both *C. botulinum* and *C. butyricum* producing BoNT/E hybridized poorly with probe NN3 (Fig. 2). The hybridization, while similar for all BoNT/E-producing strains, was much weaker than with fragments amplified from DNA of any other toxin type, as shown by the lower temperature stringency wash required (Fig. 2c). A very high degree of sequence homology has been demonstrated for the *BoNT/E* genes of *C. botulinum* and *C. butyricum*, and lateral transfer from one to the other could have occurred [11]. The high sequence homology has now been shown to extend upstream of the *BoNT* gene, including that of *NTNH* [6], indicating that this DNA was probably acquired in the same transfer event. The poor hybridization of probe NN3 with samples from type E strains may be explained by examination of the sequence of *ent*-120, which reveals three base mismatches between that of the gene and probe [7]. The nontoxic strains of *C. botulinum* group II, phenotypically identical to type E except in toxin production, showed no PCR product or hybridization signal with the internal probe (Fig. 2).

Strains of proteolytic and nonproteolytic *C. botulinum* producing BoNT/F, *C. barati* producing BoNT/F, and *C. botulinum* producing BoNT/G gave PCR products that hybridized with probe NN3 (Fig. 2). No PCR product was obtained with DNA of *Clostridium hastiforme*.

Location of *NTNH* gene. To investigate the location of the *NTNH* gene, we used oligonucleotides NN3 and

I)

a)   b)   c)
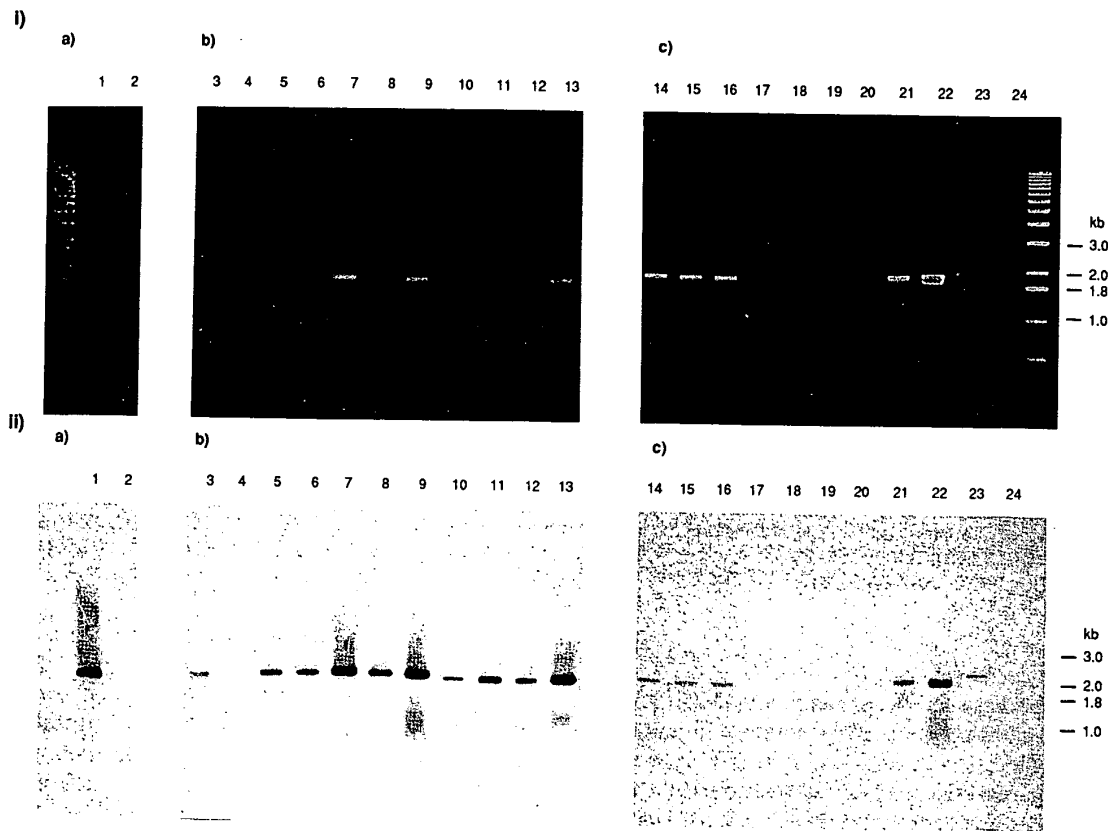
Fig. 2. (i) Agarose gel electrophoresis of PCR products obtained from a reaction with primers NN1 and NN2 and template DNA from (a) lane 1, type A, NCTC 7272; 2, type A, Kyoto-F; (b) 3, type B, NCTC 7273; 4, 'type B' nontoxic, 4672 U-1; 5, type B, Eklund 17B; 6, type B, Hobbs FT50; 7, type B, Eklund 2B; 8, type B, Colworth 151; 9, type B, Scott 2129B; 10, type F, Langeland; 11, type F, Eklund 202F; 12, type F, Hobbs FT10; 13, type F, Craig 610; (c) 14, type E, Beluga; 15, type E, Sebald P34; 16, type E, VH; 17, 'type E' nontoxic, VPI 4088; 18, 'type E' nontoxic, VPI 10428; 19, 'type E' nontoxic, VPI 2093-1; 20, 'type E' nontoxic, VPI 3158; 21, type E, C. butyricum ATCC 43755; 22, type F, C. barati ATCC 43756; 23, type G, ATCC 27322; 24, C. hastiforme DSM 5675. Marker DNA used was 1 kb ladder. (ii) Blots of the gels shown in (i) hybridized with probe NN3 and washed at 48°C for (a) and (b), and at 42°C for (c).
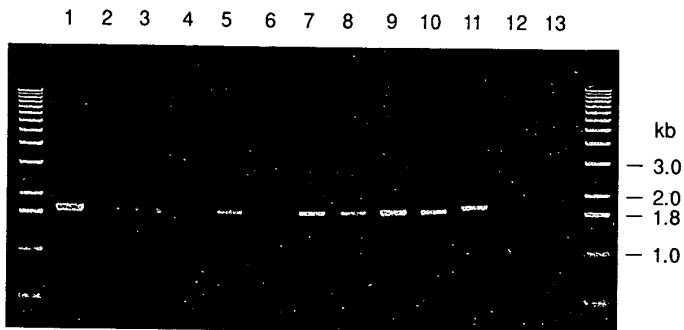
Fig. 3. Agarose gel electrophoresis of PCR products obtained from a reaction with primers NN3 and BT1 (BT2 for types B and G) and template DNA from lane 1, type A, NCTC 7272; 2, type A, Kyoto-F; 3, type B, Eklund 2B; 4, 'type B' nontoxic, 4672 U-1; 5, type E, Beluga; 6, 'type E' nontoxic, VPI 3158; 7, type E, C. butyricum ATCC 43755; 8, type F, Eklund 202F; 9, type F, Langeland; 10, type F, C. barati ATCC 43756; 11, type G, ATCC 27322; 12, C. hastiforme DSM 5675.

BT1 or BT2 (BT1 and BT2 are designed to the conserved region IWIIPER near the N-terminus of BoNT, amino acids 43 to 49 of BoNT/F, Fig. 4) as PCR primers with template DNA from representative strains of each of the immunological toxin types.

All toxin-producing strains gave PCR products of a similar size (~1.8 kb) (Fig. 3), corresponding to the fragment spanning the BoNT and NTNH genes (shown in Fig. 1(i)). These results show that for all strains producing BoNT, the gene for NTNH component of

*Sau*3Al
```
GATCTTATACCTGCAGCTATAATAAATTCAGTAGTATATTGGTCTATGGAAAGTATCAAA   60
 D   L   I   P   A   A   I   I   N   S   V   V   Y   W   S   M   E   S   I   K

GAAGCATTAGGATTCCAATTGGCAAATAATTTTACAGATATAATAAATAATATTGTTGAA  120
 E   A   L · G   F   Q   L   A   N   N   F   T   D   I   I   N   N   I   V   E

TGGAATAATATAAAAATTTCAGAAGTTACAAATGTAATATTAAATGTAGGATTTTGTATA  180
 W   N   N   I   K   I   S   E   V   T   N   V   I   L   N   V   G   F   C   I

CAAGGAAATATGAATTAAAGAAGGTGAAATATATGAAAATAAATGATGATTTAAATATAA  240
 Q   G   N   M   N   *           M   K   I   N   D   D   L   N   I   (9)
        NN1
ATTCTCCTGTAGATAATAAAAATGTTGTAATAGTTAGAGCAAGAAAGACGAATATTTTTT  300
 N   S   P   V   D   N   K   N   V   V   I   V   R   A   R   K   T   N   I   F  (29)
                                                          710
TTAAGGCGTTTCAAGTTGCTCCTAATATTTGGGTAGCTCCAGAAAGGTATTATGGAGAAC  360
 F   K   A   F   Q   V   A   P   N   I   W   V   A   P   E   R   Y   Y   G   E  (49)
                  Sau3Al
CATTGAATATAAGTGATCAAGAAAAATCTGATGGTGGAATATATGATGAGAATTTTCTTA  420
 P   L   N   I   S   D   Q   E   K   S   D   G   G   I   Y   D   E   N   F   L  (69)

AAGAAAATTCTGAAAAAGAAGAATTTTTGCAAGCAATTATTTTACTACTTAAAAGAATAA  480
 K   E   N   S   E   K   E   E   F   L   Q   A   I   I   L   L   L   K   R   I  (89)

ATAATAATATCATAGGCCAAAAGCTACTATCATTGATGTGCACATCTATTCCATTTCTTC  540
 N   N   N   I   I   G   Q   K   L   L   S   L   M   C   T   S   I   P   F   L  (109)

ATGAATACAAACAAGGAGATTATAGGCAAAGTAATTATCTTGGCTCTAAAAATAGCGAAT  600
 H   E   Y   K   Q   G   D   Y   R   Q   S   N   Y   L   G · S   K   N   S   E  (129)

ATCTCTATTCTGCAAATATAGTTATTTTTGGTCCAGGATCAAATATAGTAAAAAATAATA  660
 Y   L   Y   S   A   N   I   V   I   F   G   P   G   S   N   I   V   K   N   N  (149)

CTATTTATTATAAAAAAAATTTTGCGGAAAACGGTATGGGAACTATGGCAGAAATATTAT  720
 T   I   Y   Y   K   K   N   F   A   E   N   G   M   G   T   M   A   E   I   L  (169)

TTCAACCCCTTTTTAACTTATAAATATAATCAATTTTATGCTGATCCTGCACTAGAATTAA  780
 F   Q   P · L   L   T   Y   K   Y   N   Q   F   Y   A   D   P   A   L   E   L· (189)

TAAAATGCTTAATAAAAGCCATATATTTCTTGTATGGAATAAAACCTAATGATAATCTAA  840
 I   K   C   L   I   K   A   I   Y   F   L   Y   G   I   K   P   N · D   N   L  (209)

ATATTCCTTATAGACTAAGAAATGAATTTAGTAATGTTGAATATTCAGAATTAAATATTA  900
 N   I   P   Y   R   L   R   N   E   F   S   N   V   E   Y   S   E   L   N   I  (229)

TTGATTTTTTAATATCAGGAGGAATTGATTATAAATTTATAAATACTAATCCTTATTGGT  960
 I   D   F   L   I   S   G   G   I   D   Y   K   F   I   N   T   N   P   Y   W  (249)

TTATAGATAATTATTTTATTGACGTTCCAAAAGTTTTTGAAAAACATAAAAATGATTATG  1020
 F   I   D   N   Y   F   I   D   V   P   K   V   F   E   K   H   K   N   D   Y  (269)

AAATAAACATTAAAAATAATAGTGAAATTGGGACTAGTATTAAATTATATTTAGAGCAAA  1080
 E   I   N   I   K   N   N   S   E   I   G   T   S   I   K   L   Y   L   E   Q  (289)

AATTTAAAACTAATGTACAAGATATATGGGAACTAAATTTAAGTTATTTTTCTAAAGAAT  1140
 K   F   K   T   N   V   Q   D   I   W   E   L   N   L   S   Y   F · S   K   E  (309)

TTCAAATTATGATGCCAGAAAAACATAATAATGCACTTAAGCATTATTATAGAAAAGAAT  1200
 F   Q   I   M   M   P   E   K   H   N   N   A   L   K   H   Y   Y   R   K   E  (329)

ACTATAAAATAAATTACTCTAAACAATATGATATAAACGGCTTTGTTAATGGTCAAATTG  1260
 Y   Y   K   I   N   Y   S   K   Q   Y   D   I   N   G   F   V   N   G   Q   I  (349)

CAACAAAGTTGCTTTTATCAGAAAAAAAATCAATATATTATAAACAAACCCCAATTAATTA  1320
 A   T   K   L   L   L   S   E   K   N   Q   Y   I   I   N   K   P   Q   L   I  (369)
```

Fig. 4. Complete nucleotide sequence of the gene encoding the NTNH component of the botulinum toxin complex encoded by nonproteolytic *C. botulinum* type F (202F). The position and sequence of primers used in PCR is indicated, with mismatches between primer and the sequence shown indicated by underlining. The amino acid sequence of the protein is given under the second base of each codon.

M-complex lies directly upstream of the *BoNT* gene. The nontoxigenic strains gave no PCR product with primers NN3 and BT1 (Fig. 3).

Sequence determination of *NTNH/F* gene. The sequence of the gene encoding NTNH/F was deter-

mined after cloning fragments amplified by PCR from *C. botulinum* (strain 202F) chromosomal DNA (outlined in Fig. 1(ii)). PCR primers were designed with known sequence upstream of *BoNT/F* [4] and based on regions of *NTNH/C* [18] that were considered likely to be conserved. Inverse PCR was used to

```
TTAATTTAATAAATAAAAGTAATAATAGTTTATTAATGAAAAGTAACATTTATGGAGATG   1380
I  N  L  I  N  K  S  N  N  S  L  L  M  K  S  N  I  Y  G  D     (389)

GATTAAATGGAACTACAGATAATTTTTACAGGAATTATAAAATTCCAGATAATATAGCTT   1440
G  L  N  G  T  T  D  N  F  Y  R  N  Y  K  I  P  D  N  I  A     (409)

ATCAATATCATCCTAATAATACTTATTTAGATAATGTAAATATTGAAGAAATAAATAATA   1500
Y  Q  Y  H  P  N  N  T  Y  L  D  N  V  N  I  E  E  I  N  N     (429)

TTCCTCAAATTACTGATGCAGATATATATCCATATACAAATAATTGTGATACTTTTATAC   1560
I  P  Q  I  T  D  A  D  I  Y  P  Y  T  N  N  C  D  T  F  I     (449)

CTATATATAATATAACTCAATCTAGAGAAATAAATACTACTGTACCTTATTCTATAAATT   1620
P  I  Y  N  I  T  Q  S  R  E  I  N  T  T  V  P  Y  S  I  N     (469)

ATTTGCAGTCTCAAATTATGAATAGTGACGATATTACTCTATCTTCTGATTTTTGGGAAG   1680
Y  L  Q  S  Q  I  M  N  S  D  D  I  T  L  S  S  D  F  W  E     (489)

TAGTTTGTTCAAATGATAAATCTTTAGTATATTCTTATTTAGATAATGTAATAAATTATT   1740
V  V  C  S  N  D  K  S  L  V  Y  S  Y  L  D  N  V  I  N  Y     (509)

TAGATTCTATAAAAAATAATACTCCAATTAATACAGATAAAAAGTATTACTTATGGCTAA   1800
L  D  S  I  K  N  N  T  P  I  N  T  D  K  K  Y  Y  L  W  L     (529)

AAGAAATATTTAGGAACTATTCTTTTGATATTACAGCAACTGAAGAAATTACTACTGAAT   1860
K  E  I  F  R  N  Y  S  F  D  I  T  A  T  E  E  I  T  T  E     (549)

GTGGAATTAATAAAATAGTATCTTGGTTTGGAAAAGCTTTAAATATTTTAAATACAGATA   1920
C  G  I  N  K  I  V  S  W  F  G  K  A  L  N  I  L  N  T  D     (569)

ATTCTTTTAAAATAGAATTTCAAAATTCTGGGGCAATTGCTCTTATTAATAAAAAAGACA   1980
N  S  F  K  I  E  F  Q  N  S  G  A  I  A  L  I  N  K  K  D     (589)

ATATAATCATTCCAAAAATAGAGATTGATGAAATGCCTAATAGTATGTTAAATTTATCAT   2040
N  I  I  I  P  K  I  E  I  D  E  M  P  N  S  M  L  N  L  S     (609)

TTGAAGATTTAAATGAACAGTTATATAGTATATATTCTAAAAATATAACTTACTTTAAAA   2100
F  E  D  L  N  E  Q  L  Y  S  I  Y  S  K  N  I  T  Y  F  K     (629)
                     NN3
                 ──────────────────────────►
AAATCTATTATAATTTCTTAGATCAATGGTGGACAGAATATTATAGTCAATATTTTGATT   2160
K  I  Y  Y  N  F  L  D  Q  W  W  T  E  Y  Y  S  Q  Y  F  D     (649)
                              NN2
                   ◄──────────────────────
TAATTTGTATGGCAAAAAAATCAATATTAGCACAAGAAAATTTAATCAAAAAAATAATAC   2220
L  I  C  M  A  K  K  S  I  L  A  Q  E  N  L  I  K  K  I  I     (669)

AAAAGAAAATAAGCTATTTAATTGGGGCATCTAATATACCTGATGATATATTAGCAGTAA   2280
Q  K  K  I  S  Y  L  I  G  A  S  N  I  P  D  D  I  L  A  V     (689)

TGAGATTGACAACAACTAATACTTTAAGAGATATATCAGTCGAATCACAAATAGCAATGA   2340
M  R  L  T  T  T  N  T  L  R  D  I  S  V  E  S  Q  I  A  M     (709)

ATAATTTAAATAATTTTTTAAATAAAGCGGCTATGTGTGTGTTTCAAAGTAATATTTATC   2400
N  N  L  N  N  F  L  N  K  A  A  M  C  V  F  Q  S  N  I  Y     (729)

CTAAATTTATCTCTTTCATGGAACAATGTATTAAGCATATAAACAAAAGCACTAAAGAGT   2460
P  K  F  I  S  F  M  E  Q  C  I  K  H  I  N  K  S  T  K  E     (749)

TTATACAAAAATGTACTAATATTAATGAAACTGAAAAATTACAATTGATTATGCAAAATA   2520
F  I  Q  K  C  T  N  I  N  E  T  E  K  L  Q  L  I  M  Q  N     (769)

GTTTTAGTAATTTAGATTTTGATTTTTTAGATATTCAAAATATGAAGAATCTATTTAACT   2580
S  F  S  N  L  D  F  D  F  L  D  I  Q  N  M  K  N  L  F  N     (789)

CATATACAGAACTACTTATAAAAGAACAAACCTCACCCTATGAATTATCATTATATGCTT   2640
S  Y  T  E  L  L  I  K  E  Q  T  S  P  Y  E  L  S  L  Y  A     (809)
```
Fig. 4. Continued.

obtain the extreme 5'-end of the gene and upstream sequences.

The complete nucleotide sequence of the *NTNH/F* gene is given in Fig. 4. The gene encodes a protein of 1165 amino acids with a calculated rmm of 136,525. The *NTNH* gene ends 14 nucleotides before the start of *BoNT/F*. For other BoNT types the distance between these two genes varies, e.g., 43 nucleotides for type A [1]. Directly upstream of the start of *NTNH/F* gene there are no recognized −10 and −35 sequences. Further experiments need to be performed before conclusions can be drawn on what sequences form the promoter, the start of the mRNA, and whether the message is mono- or poly-cistronic.

```
TTGAAGAACAAGATAACAATGTTATAGGAGATGCATCAGGTAAAAATACATTAGTAGAAT   2700
F   E   E   Q   D   N   N   V   I   G   D   A   S   G   K   N   T   L   V   E    (829)

ACCCTAAAGGAATAGAATTAGTTTATGGAATAAATAATAGTGCATTATATTTAAATGGAT   2760
Y   P   K   G   I   E   L   V   Y.-G   I   N   N   S   A   L   Y   L   N   G    (849)

CTAATCAAAGTATAATTTTTACCAACGATTATTTTGAGAATGGACTAACTAATAGTTTTT   2820
S   N   Q   S   I   I   F   T   N   D   Y   F   E   N   G   L   T   N   S   F'   (869)

CAATTTATTTTTGGTTGAGAAATTTGGGCCAAGATACTATTAAATCTAAGTTAATAGGTA   2880
S   I   Y   F   W   L   R   N   L   G   Q   D   T   I   K   S   K   L   I   G    (889)

GTAAAGAATATAATTGTGGTTGGGAAATTTATTTTCAAGAGATTGGGCATGTTTTTAATA   2940
S   K   E   Y   N   C   G   W   E   I   Y   F   Q   E   I   G   H   V   F   N    (909)

TGATAGATTCTAATGGAAATGAGAAAAATATATATCTATCCGATGTTTCTAATAATAGTT   3000
M   I   D   S   N   G   N   E   K   N   I   Y   L   S   D   V   S   N   N   S    (929)
                              Sau3AI
GGCACTATATAACTATATCTGTGGATCGTTTAAAAGAACAATTATTAATATTTATTGATG   3060
W   H   Y   I   T   I   S   V   D   R   L   K   E   Q   L   L   I   F   I   D    (949)

ATAATTTAGTAGTTAATGAAAGTATTAAAGATATTTTAAATATTTATTCAAGTAATATAA   3120
D   N   L   V   V   N   E   S   I.  K   D   I   L   N   I   Y   S   S   N   I    (969)
                           ◄————  657  ————————————
TTTCTTTATTAAGTGATAATAAGGCAAGTTATATTGAGGGATTAACTATTTTAAATAAAC   3180
I   S   L   L   S   D   N   K   A   S   Y   I   E   G   L   T   I   L   N   K    (989)

CTACTACAGGTGAAGAAGTTTTGAGAAATTATTTTAAGAACCTAAATAATTCATATGTAA   3240
P   T   T   G   E   E   V   L   R   N   Y   F   K   N   L   N   N   S   Y   V    (1009)

GAGATAGTAATGATGAACGATTAGAATATAATAAGACATATCAATTATATGATTATGTAT   3300
R   D   S   N   D   E   R   L   E   Y   N   K   T   Y   Q   L   Y.  D   Y   V    (1029)

TTCCAGATAATCCTATATGTGAAGTTAAACAAGATAATAATATATATTTAACAATTAATA   3360
F   P   D   N   P   I   C   E   V   K   Q   D   N   N   I   Y   L   T   I   N    (1049)

ATATAAATAATTTAAATATGAAACCTTGTAAATTTAAATTATTAAGTATTAATTCAAATA   3420
N   I   N   N'  L   N   M   K   P   C   K   F   K   L   L   S   I   N   S   N    (1069)

AACAATATGTTCAGAAATGGGATGAAGTAATAATATCTGTATTATATGATACAGAAAAAT   3480
K   Q   Y   V   Q   K   W   D   E   V   I   I   S   V   L   Y   D   T   E   K    (1089)

ATGTATGTATATCTAATGAAAATAATAGAGTTAAAATAATAGATAATAAAATAATGCAAG   3540
Y   V   C   I   S   N   E   N   N   R   V   K   I   I   D   N   K   I   M   Q    (1109)

TAAAATTTATTATTAGTAATGATATATTTATTTCAAATTGTTTAACGCATGCTCATAACA   3600
V   K   F   I   I   S   N   D   I   F   I   S   N   C   L   T   H   A   H   N    (1129)

ATAAGTATATATGTTTATCTATGAAAGATGAGAACTATAATTGGATGATATGTAACAATG   3660
N   K   Y   I   C   L   S   M   K   D   E   N   Y   N   W   M   I   C   N   N    (1149)

AAAGCAACATACCTAAAAAGGCATATTTGTGGATATTGAAAGAAGTATAGGGGGGATTT   3720
E   S   N   I   P   K   K   A   Y   L   W   I   L   K   E   V   *    (1165)

TATGCCAGTTGCAATAAATAGTTTTAATTATAATGACCCTGTTAATGATGATACAATTTT   3780
    M   P   V   A   I   N   S   F   N   Y   N   D   P   V   N   D   D   T   I   L    (20)
         ◄————  236  ————————————                              164  ————
ATACATGCAGATACCATATGAAGAAAAAAGTAAAAAATATTATAAAGCTTTTGAGATTAT   3840
    Y   M   Q   I   P   Y   E   E   K   S   K   K   Y   Y   K   A   F   E   I   M    (40)
              ◄————————————————  BT1
              ◄————————————  BT2                                   Sau3AI
GCGTAATGTTTGGATAATTCCTGAGAGAAATACAATAGGAACGAATCCTAGTGATTTTGA   3900
    R   N   V   W   I   I   P   E   R   N   T   I   G   T   N   P   S   D   F   D    (60)

TCCACCGGCTTCA  ————►  BoNT/F                                     3913
    P   P   A   S                                                  (64)   Fig. 4. Continued.
```

Upstream of the genes encoding NTNH/F, Ent-120 (NTNH/E), and NTNH of *C. butyricum* there are sequences which could be the 3'-end of an open reading frame [Fig. 4]. These nonproteolytic strains of *C. botulinum* do not produce hemagglutinin, the gene for which is found immediately upstream of *NTNH/C* [18], and the sequence encoded here shows no homology to the hemagglutinin of type C [17]. Since type E and F DNA sequences are homologous in this region upstream of genes encoding the components of botulinum protein complex M, the sequences could also be concerned with the BoNT-complex,
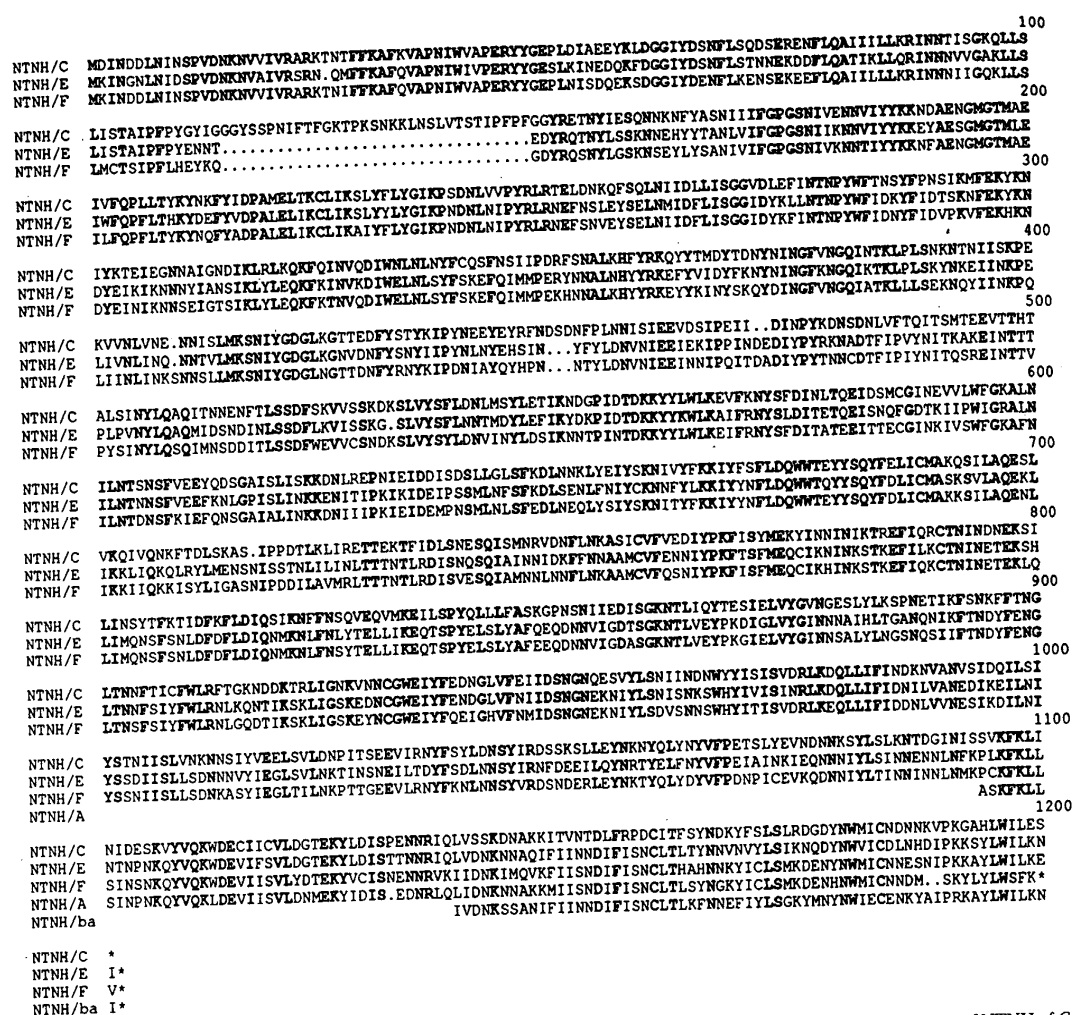
76

```
                                                                              100
NTNH/C   MDINDDLNINSPVDNKNVVIVRARKTNTFFKAFKVAPNIWVAPERYYGEPLDIAEEYKLDGGIYDSNFLSQDSERENFLQAIIILLKRINNTISGKQLLS
NTNH/E   MKINGNLNIDSPVDNKNVAIVRSRN.QMFFKAFQVAPNIWIVPERYYGESLKINEDQKFDGGIYDSNFLSTNNEKDDFLQATIKLLQRINNNVVGAKLLS
NTNH/F   MKINDDLNINSPVDNKNVVIVRARKTNIFFKAFQVAPNIWVAPERYYGEPLNISDQEKSDGGIYDENFLKENSEKEEFLQAIILLLKRINNNIIGQKLLS
                                                                              200
NTNH/C   .LISTAIPFPYGYIGGGYSSPNIFTFGKTPKSNKKLNSLVTSTIPFPFGGYRETNYIESQNNKNFYASNIIIFGPGSNIVENNVIYYKKNDAENGMGTMAE
NTNH/E   LISTAIPFPYENNT...............................EDYRQTNYLSSKNNEHYYTANLVIFGPGSNIIKNNVIYYKKEYAESGMGTMLE
NTNH/F   LMCTSIPFLHEYKQ...............................GDYRQSNYLGSKNSEYLYSANIVIFGPGSNIVKNNTIYYKKNFAENGMGTMAE
                                                                              300
NTNH/C   IVFQPLLTYKYNKFYIDPAMELTKCLIKSLYFLYGIKPSDNLVVPYRLRTELDNKQFSQLNIIDLLISGGVDLEFINTNPYWFTNSYFPNSIKMFEKYKN
NTNH/E   IWFQPFLTHKYDEFYVDPALELIKCLIKSLYYLYGIKPNDNLNIPYRLRNEFNSLEYSELNMIDFLISGGIDYKLLNTNPYWFIDKYFIDTSKNFEKYKN
NTNH/F   ILFQPFLTYKYNQFYADPALELIKCLIKAIYFLYGIKPNDNLNIPYRLRNEFSNVEYSELNIIDFLISGGIDYKFINTNPYWFIDNYFIDVPKVFEKHKN
                                                                              400
NTNH/C   IYKTEIEGNNAIGNDIKLRLKQKFQINVQDIWNLNLNYFCQSFNSIIPDRFSNALKHFYRKQYYTMDYTDNYNINGFVNGQINTKLPLSNKNTNIISKPE
NTNH/E   DYEIKIKNNNYIANSIKLYLEQKFKINVKDIWELNLSYFSKEFQIMMPERYNNALNHYYRKEFYVIDYFKNYNINGFKNGQIKTKLPLSKYNKEIINKPE
NTNH/F   DYEINIKNNSEIGTSIKLYLEQKFKTNVQDIWELNLSYFSKEFQIMMPEKHNNALKHYYRKEYYKINYSKQYDINGFVNGQIATKLLLSEKNQYIINKPQ
                                                                              500
NTNH/C   KVVNLVNE.NNISLMKSNIYGDGLKGTTEDFYSTYKIPYNEEYEYRFNDSDNFPLNNISIEEVDSIPEII..DINPYKDNSDNLVFTQITSMTEEVTTHT
NTNH/E   LIVNLINQ.NNTVLMKSNIYGDGLKGNVDNFYSNYIIPYNLNYEHSIN...YFYLDNVNIEEIEKIPPINDEDIYPYRKNADTFIPVYNITKAKEINTTT
NTNH/F   LIINLINKSNNSLLMKSNIYGDGLNGTTDNFYRNYKIPDNIAYQYHPN...NTYLDNVNIEEINNIPQITDADIYPYTNNCDTFIPIYNITQSREINTTV
                                                                              600
NTNH/C   ALSINYLQAQITNNENFTLSSDFSKVVSSKDKSLVYSFLDNLMSYLETIKNDGPIDTDKKYYLWLKEVFKNYSFDINLTQEIDSMCGINEVVLWFGKALN
NTNH/E   PLPVNYLQAQMIDSNDINLSSDFLKVISSKG.SLVYSFLNNTMDYLEFIKYDKPIDTDKKYYKWLKAIFRNYSLDITETQEISNQFGDTKIIPWIGRALN
NTNH/F   PYSINYLQSQIMNSDDITLSSDFWEVVCSNDKSLVYSYLDNVINYLDSIKNNTPINTDKKYYLWLKEIFRNYSFDITATEEITTECGINKIVSWFGKAFN
                                                                              700
NTNH/C   ILNTSNSFVEEYQDSGAISLISKKDNLREPNIEIDDISDSLLGLSFKDLNNKLYEIYSKNIVYFKKIYFSFLDQWWTEYYSQYFELICMAKQSILAQESL
NTNH/E   ILNTNNSFVEEFKNLGPISLINKKENITIPKIKIDEIPSSMLNFSFKDLSENLFNIYCKNNFYLKKIYYNFLDQWWTQYYSQYFDLICMASKSVLAQEKL
NTNH/F   ILNTDNSFKIEFQNSGAIALINKKDNIIIPKIEIDEMPNSMLNLSFEDLNEQLYSIYSKNITYFKRIYYNFLDQWWTEYYSQYFDLICMARKSILAQENL
                                                                              800
NTNH/C   VKQIVQNKFTDLSKAS.IPPDTLKLIRETTEKTFIDLSNESQISMNRVDNFLNKASICVFVEDIYPKFISYMEKYINNINIKTREFIQRCTNINDNEKSI
NTNH/E   IKKLIQKQLRYLMENSNISSTNLILINLTTTNTLRDISNQSQIAINNIDKFFNNAAMCVFENNIYPKFTSFMEQCIKNINKSTKEFILKCTNINETEKSH
NTNH/F   IKKIIQKKISYLIGASNIPDDILAVMRLTTTNTLRDISVESQIAMNNLNNFLNKAAMCVFQSNIYPKFISFMEQCIKHINKSTKEFIQKCTNINETEKLQ
                                                                              900
NTNH/C   LINSYTFKTIDFKFLDIQSIKNFFNSQVEQVMKEILSPYQLLLFASKGPNSNIIEDISGKNTLIQYTESIELVYGVNGESLYLKSPNETIKFSNKFFTNG
NTNH/E   LIMQNSFSNLDFDFLDIQNMKNLFNLYTELLIKEQTSPYELSLYAFQEQDNNVIGDTSGKNTLVEYPKDIGLVYGINNNAIHLTGANQNIKFTNDYFENG
NTNH/F   LIMQNSFSNLDFDFLDIQNMKNLFNSYTELLIKEQTSPYELSLYAFEEQDNNVIGDASGKNTLVEYPKGIELVYGINNSALYLNGSNQSIIFTNDYFENG
                                                                              1000
NTNH/C   LTNNFTICFWLRFTGKNDDKTRLIGNKVNNCGWEIYFEDNGLVFEIIDSNGNQESVYLSNIINDNWYYISISVDRLKDQLLIFINDKNVANVSIDQILSI
NTNH/E   LTNNFSIYFWLRNLKQNTIKSKLIGSKEDNCGWEIYFENDGLVFNIIDSNGNEKNIYLSNISNKSWHYIVISINRLKDQLLIFIDNILVANEDIKEILNI
NTNH/F   LTNSFSIYFWLRNLGQDTIKSKLIGSKEYNCGWEIYFQEIGHVFNMIDSNGNEKNIYLSDVSNNSWHYITISVDRLKEQLLIFIDDNLVVNESIKDILNI
                                                                              1100
NTNH/C   YSTNIISLVNKNNSIYVEELSVLDNPITSEEVIRNYFSYLDNSYIRDSSKSLLEYNKNYQLYNYVFPETSLYEVNDNNKSYLSLKNTDGINISSVKFKLI
NTNH/E   YSSDIISLLSDNNNVYIEGLSVLNKTINSNEILTDYFSDLNNSYIRNFDEEILQYNRTYELFNYVFPEIAINKIEQNNNIYLSINNENNLNFKPLKFKLL
NTNH/F   YSSNIISLLSDNKASYIEGLTILNKPTTGEEVLRNYFKNLNNSYVRDSNDERLEYNKTYQLYDYVFPDNPICEVKQDNNIYLTINNINNLNMKPCKFKLL
NTNH/A                                                                   ASKFKLL
                                                                              1200
NTNH/C   NIDESKVYVQKWDECIICVLDGTEKYLDISPENNRIQLVSSKDNAKKITVNTDLFRPDCITFSYNDKYFSLSLRDGDYNWMICNDNNKVPKGAHLWILES
NTNH/E   NTNPNKQYVQKWDEVIFSVLDGTEKYLDISTTNNRIQLVDNKNNAQIFIINNDIFISNCLTLTYNNVNVYLSIKNQDYNWVICDLNHDIPKKSYLWILKN
NTNH/F   SINSNKQYVQKWDEVIISVLYDTEKYVCISNENNRVKIIDNKIMQVKFIISNDIFISNCLTHAHNNKYICLSMKDENYNWMICNNESNIPKKAYLWILKE
NTNH/A   SINPNKQYVQKLDEVIISVLDNMEKYIDIS.EDNRLQLIDNKNNAKKMIISNDIFISNCLTLSYNGKYICLSMKDENHNWMICNNDM..SKYLYLWSFK*
NTNH/ba                        IVDNKSSANIFIINNDIFISNCLTLKFNNEFIYLSGKYMNYNWIECENKYAIPRKAYLWILKN

NTNH/C   *
NTNH/E   I*
NTNH/F   V*
NTNH/ba  I*
```

Fig. 5. Multiple alignment of amino acid sequences of NTNH/C, NTNH/E, and NTNH/F, and partial amino acid sequences of NTNH of C. barati [16] and C. botulinum type A (translated from [1]). Amino acids that are identical in all proteins are shown in bold.

perhaps encoding proteins involved in formation of the complex or control of expression of the constituent proteins.

Comparison of NTNH/F with other NTNHs. A multiple alignment of NTNH/F, NTNH/C, and NTNH/E is shown in Fig. 5. The NTNH/F shows 71% identity with NTNH/E and 61% identity with NTNH/C (Table 2). For types C, E, and F the NTNH component of the toxin complex is more highly conserved than BoNT (Table 2). This is perhaps surprising when the proposed function of the two proteins is considered; BoNT is a highly specific and potent neurotoxin, while the suggested function of NTNH is to protect the toxin from proteolytic degradation [15].

NTNH/F and NTNH/E have a common deletion

Table 2. Percentage similarities (lower left-hand triangle) and identities (upper right-hand triangle) of (i) NTNH and (ii) BoNT amino acid sequences

(i) NTNH				(ii) BoNT			
Type	C	E	F	Type	C	E	F
C	—	57.9	60.7	C	—	34.5	34.5
E	75.0	—	71.2	E	56.0	—	62.6
F	76.1	83.5	—	F	56.2	77.4	—

with respect to NTNH/C (Fig. 5). From PCR results it seems that only the NTNH components from types E and F have this deletion, as PCR products from types A, B, and G are slightly larger (Fig. 2). A comparison of BoNT sequences showed that BoNT/E

and BoNT/F are more closely related to each other than to other toxin types [8]. The high sequence homology of the NTNH components may be significant in considering the evolution of the botulinum toxin 'operon'.

The results presented here show that all BoNT-producing strains tested possess a gene encoding NTNH directly upstream of the toxin gene, showing that the genes of the two components of the M-complex are structurally linked. Since only some of the BoNT-producing strains make the larger (L- and LL-) forms of the progenitor complex, it will be interesting to compare the structural maps of genes encoding components of the complexes in strains producing L- and LL-forms with those that produce only the M-complex.

ACKNOWLEDGMENTS

Literature Cited

1. Binz T, Kurazono H, Wille M, Frevert J, Wernars K, Niemann H (1990) The complete sequence of botulinum neurotoxin type A and comparison with other clostridial neurotoxins. J Biol Chem 265:9153–9158
2. Campbell KD, Collins MD, East AK (1993) Gene probes for identification of the botulinal neurotoxin gene and specific identification of neurotoxin types B, E, and F. J Clin Microbiol 31:2255–2262
3. Devereux J, Haeberli P, Smithies O (1984) A comprehensive set of sequence analysis programs for the VAX. Nucleic Acids Res 12:387–395
4. East AK, Richardson PT, Allaway D, Collins MD, Roberts TA, Thompson DE (1992) Sequence of the gene encoding type F neurotoxin of Clostridium botulinum. FEMS Microbiol Lett 96:225–230
5. Fujii N, Kimura K, Yokosawa N, Yashiki T, Tsuzuki K, Oguma K (1993a) The complete nucleotide sequence of the gene encoding the nontoxic component of Clostridium botulinum type E progenitor toxin. J Gen Microbiol 139:79–86
6. Fujii N, Kimura K, Yokosawa N, Oguma K, Yashiki T, Takeshi K, Ohyama T, Isogai E, Isogai H (1993b) Similarity in nucleotide sequence of the gene encoding nontoxic component of botulinum toxin produced by toxigenic Clostridium butyricum strain BL6340 and Clostidium botulinum type E strain Mashike. Microbiol Immunol 37:395–398
7. Hatheway CL (1992) Clostridium botulinum and other clostridia that produce botulinum neurotoxin. In: Hauschild AHW, Dodds KL (eds) Clostridium botulinum—ecology and control in foods. New York: Marcel Dekker Inc., pp. 3–20
8. Hutson RA, Collins MD, East AK, Thompson DE (1993) Nucleotide sequence of the gene coding for non-proteolytic type B neurotoxin: comparison with other clostridial toxins. Curr Microbiol, in press.
9. McCroskey LM, Hatheway CL, Fenicia L, Pasolini B, Aureli P (1986) Characterization of an organism that produces type E botulinal toxin but which resembles Clostridium butyricum from the feces of an infant with type E botulism. J Clin Microbiol 23:201–202
10. McCroskey LM, Hatheway CL, Woodruff BA, Greenberg JA, Jurgenson PA (1991) Type F botulism due to neurotoxigenic Clostridium baratii from an unknown soure in an adult. J Clin Microbiol 29:2618–2620
11. Poulet S, Hauser D, Quanz M, Niemann H, Popoff MR (1992) Sequences of the botulinal neurotoxin E derived from Clostridium botulinum type E (strain Beluga) and Clostridium butyricum (strains ATCC 43181 and ATCC 43755). Biochem Biophys Res Commun 183:107–113
12. Sakaguchi G, Sakaguchi S, Kamat Y, Tabita K, Asao T, Kozaki S (1990) Distinct characters of Clostridium botulinum type A strains and their toxin associated with infant botulism in Japan. Int J Food Microbiol 11:231–242
13. Sambrook J, Fritsch EF, Maniatis T (1989) Molecular cloning, a laboratory manual, 2nd ed. Cold Spring Harbor, N.Y.: Cold Spring Harbor Laboratory Press
14. Somers E, DasGupta BR (1991) Clostridium botulinum types A, B, C1 and E produce proteins with or without hemagglutinating activity: do they share common amino acid sequences and genes? J Prot Chem 10:415–425
15. Tabita K, Sakaguchi S, Kozaki S, Sakaguchi G (1991) Distinction between Clostridium botulinum type A strains associated with food-borne botulism and those with infant botulism in Japan in intraintestinal toxin production in infant mice and some other properties. FEMS Microbiol Lett 79:251–256
16. Thompson DE, Hutson RA, East AK, Allaway D, Collins MD, Richardson PT (1993) Nucleotide sequence of the gene coding for Clostridium barati type F neurotoxin: comparison with other clostridial neurotoxins. FEMS Microbiol Lett 108:175–182
17. Tsuzuki K, Kimura K, Fuji N, Yokosawa N, Indoh T, Murakami T, Oguma K (1990) Cloning and complete nucleotide sequence of the gene for the main component of hemagglutinin produced by Clostridium botulinum type C. Infect Immun 58:3173–3177
18. Tsuzuki K, Kimura K, Fuji N, Yokosawa N, Oguma K (1992) The complete nucleotide sequence of the gene coding for the nontoxic-nonhemagglutinin component of Clostridium botulinum type C progenitor toxin. Biochem Biophys Res Commun 183:1273–1279

Hc-peptides c rresponding to these two regions were synthesized and were demonstrated to bind the neutralizing mAbs. Mice immunized with the Hc-peptides had high levels of antibodies that recognized BoNT/A-Hc. H wever, immunizations with only one of the Hc peptides protected when mice were challenged with BoNT/A. On the basis of these analyses, it should be possible to develop small peptides that could be useful in the design of future vaccines against these neurotoxins.

Thanks,
Sharon L. Turner, Ph.D.
USPTO
CM1-10B09
Mailroom 10C01
Biotechnology GAU 1647
(703) 308-0056

AMINO ACID COMPOSITION OF *CLOSTRIDIUM BOTULINUM* TYPE F NEUROTOXIN

BIBHUTI R. DASGUPTA and STUART RASMUSSEN

Food Research Institute, University of Wisconsin, Madison, WI 53706, U.S.A.

B. R. DASGUPTA and S. RASMUSSEN. Amino acid composition of *Clostridium botulinum* type F neurotoxin. *Toxicon* 21, 566–569, 1983. — To develop reliable data on the amino acid composition of type F botulinum neurotoxin, three batches of the neurotoxin were analyzed. Each batch was isolated from a separate neurotoxin producing bacterial culture. Two batches had inoculum from one source and the other batch one from a different source. Two batches of the neurotoxin were purified by the same method and one was purified by a different method. The neurotoxin preparations were found comparable in purity and similar in amino acid composition. The best estimate of number of amino acid residues per neurotoxin molecule (mol. wt. 155,000) was: Asp_{118} Thr_{80} Ser_{105} Glu_{118} Pro_{47} Gly_{69} Ala_{47} Val_{71} CyS_9 Met_{14} Ile_{128} Leu_{104} Tyr_{86} Phe_{60} Lys_{90} His_{13} Arg_{51} Trp_{23}.

THE SOLE cause for the neuroparalytic food-borne disease botulism is botulinum neurotoxin, a protein produced by *Clostridium botulinum*. The neurotoxin is found in several antigenically distinguishable forms; types A, B, C_1, D, E, F (SUGIYAMA, 1980) and the recently discovered type G (LEWIS *et al.*, 1981). The type F botulinum neurotoxin is synthesized, like types A, B, C, D and E, as a single chain protein (mol. wt. $\sim$ 150,000) and depending on the physiology of the anaerobic bacteria is cleaved (nicked) endogenously in the bacterial culture into a dichain (nicked) molecule. The two chains (mol. wt. $\sim$100,000 and $\sim$50,000) of the nicked molecule are held together by at least one disulfide bond. Nothing more is known about the structure of type F neurotoxin, although studies on the structure and structure – function relationship of type A, B, C, and E have progressed (DASGUPTA, 1981; DASGUPTA and SUGIYAMA, 1977*a*; SYUTO and KUBO, 1981). Here we report on the amino acid composition of type F neurotoxin isolated by two different methods (YANG and SUGIYAMA, 1975; OHISHI and SAKAGUCHI, 1975).

Type F neurotoxin was purified in three separate batches. Stock cultures of *Clostridium botulinum* type F (strain, Langeland) were: (i) from the Food Research Institute (used in a previous study, DASGUPTA and SUGIYAMA, 1977*b*), maintained in the toxin production medium at 8°C and used to produce batches #1 and #3 of neurotoxin; (ii) kindly supplied by Dr. Lynn S. Siegel, U.S. Army Medical Research Institute of Infectious Diseases, Frederick, MD, in cooked meat medium and stored at 8°C, which was used for batch #2. Each batch of neurotoxin was produced in 15 liters of medium of identical composition (YANG and SUGIYAMA, 1975) and incubated for 5 days at 30°C following inoculation with 1.0 ml of a stock culture. Batches #1 and #2 were purified by one method (YANG and SUGIYAMA, 1975) and batch #3 by another method (OHISHI and SAKAGUCHI, 1975). Purity of the neurotoxin preparations was assayed (see discussion in DASGUPTA and SUGIYAMA, 1977*a*) by electrophoresis in 5% polyacrylamide gels in the presence of 0.1% sodium dodecyl sulfate (PAGE-SDS) (WEBER and OSBORN, 1969).

For amino acid analysis, procedures described by MOORE and STEIN (1963) and MOORE (1972) were used. Cysteine and cystine were determined as 1/2 cystine (i.e. cysteic acid) following performic acid oxidation and 24 hr hydrolysis in 6 N HCl. Tryptophan was determined from the absorbance of the protein dissolved in 6 M guanidine-HCl (EDELHOCH, 1967). The Durrum D-500 automatic amino acid analyzer was calibrated with standard amino acids (10 nmoles each) immediately before and after analyzing the hydrolyzed protein samples. The exact detaiis of the above procedures followed and calculation of the number of amino acid residues have been described (DASGUPTA and RASMUSSEN, 1983). Molecular weight of the neurotoxin, based on the PAGE-SDS technique, was reported to be 155,000 (YANG and SUGIYAMA, 1975) and 157,000 (OHISHI and SAKAGUCHI, 1975). We used the lower value to calculate the number of amino acid residues.

Amino acids recovered from one of the three batches of the neurotoxin (batch 3) following 24, 48 and 72 hr HCl hydrolysis, performic acid oxidation – HCl hydrolysis, and the calculated ratio of tryptophan to tyrosine are tabulated in Table 1. No unusual peak was found on any of the chromatograms. The calculated number of amino acid residues per neurotoxin molecule (mol. wt. 155,000) from batches 1, 2 and 3 are given in Table 2. Agreement in amino acid content among the three batches was as follows (the error on the Durrum D-500 analyzer is ± 3%): nine amino acids (Asp, Pro, 1/2 Cys, Met, Ile, Leu, Phe, Lys, Arg) were < 3%; six amino acids (Thr, Ser, Glu, Ala, Tyr, Trp) were > 3% but < 5%; three amino acids (Gly, His, Val) were > 5%. In those cases where one batch

TABLE 1. AMINO ACID ANALYSIS OF BOTULINUM NEUROTOXIN TYPE F (BATCH #3)

	HCl hydrolysis					Minimum number of residues (His = 1)
Amino acid	Performic acid		24 hr	48 hr	72 hr	Mean (Lys = 100)	
Aspartic acid	70.24	85.92	42.44	29.79	29.37	240.36	17.66
Threonine	23.21	37.09	15.15	10.42	9.90	90.05*	6.62
Serine	26.85	34.18	19.69	13.56	12.02	116.7*	8.57
Glutamic acid	38.18	46.37	22.76	16.12	15.95	129.83	9.54
Proline	14.64	17.96	8.78	6.09	6.27	50.06	3.68
Glycine	20.67	24.24	11.60	8.25	8.10	66.18	4.86
Alanine	15.46	18.76	9.16	6.44	6.29	51.77	3.80
Valine	17.53	21.72	11.37	8.60	9.00†	73.11†	5.37
Cystine	0.40	0.26	0.42	0.22	0.13	—	—
Methionine	0.04	—	2.70	1.89	1.80	15.09	1.11
Isoleucine	35.89	43.46	22.19	15.46	17.06†	138.59†	10.18
Leucine	34.09	42.02	19.52	12.89	13.79	111.94‡	8.22
Tyrosine	0.04	0.03	16.96	12.05	12.14	97.53	7.17
Phenylalanine	18.53	22.27	11.55	8.12	8.14	65.85	4.84
Ammonia		—	—	—		
Lysine	28.64	35.23	17.45	12.45	12.31	100	7.35
Histidine	0.69	0.58	2.32	1.71	1.70	13.61	1.00
Arginine	16.08	19.98	9.97	6.96	6.90	56.36	4.14
½ Cystine	2.77	3.72	—	—	—	10.12§	0.74
Tryptophan	—	—	—	—	—	26.24¶	1.93

Columns under performic acid, 24, 48 and 72 hr: nmoles of amino acids recovered from 40 μl hydrolysate.
* Thr and Ser extrapolated to zero time (Ser 72 hr value not included).
† Val 72 hr value (at 24 and 48 hr, 65.16 and 69.08, respectively) slow release. Ile 72 hr value (at 24 and 48 hr, 127.16 and 124.18, respectively) slow release.
‡ Leu mean of 24 and 72 hr values, 111.86 and 112.02, respectively (48 hr value, 103.53, not included).
§ ½ Cys mean of two values, 9.67 and 10.56, from duplicate runs.
¶ Based on tryptophan: tyrosine = 0.269 (from batch 1).

TABLE 2. NUMBER OF AMINO ACID RESIDUES PER TYPE F BOTULINUM NEUROTOXIN (MOL. WT. = 155,000)

Amino acid	Batch 1	Batch 2	Batch 3	Mean f three	Best of three
Aspartic acid	218	218	221	219	218
Threonine	77	80	83	80	80
Serine	104	114	107	108	105
Glutamic acid	128	129	120	126	128
Proline	48	47	46	47	47
Glycine	71	68	61	67	69
Alanine	45	47	48	47	47
Valine	74	71	67	71	72
½ Cystine	9*	9*	9	9	9
Methionine	14	14	14	14	14
Isoleucine	128	124	128	127	128
Leucine	104	104	103	104	104
Tyrosine	87	86	90	87	86
Phenylalanine	60	59	61	60	60
Lysine	90	90	92	91	90
Histidine	10	13	13	12	13
Arginine	51	51	52	51	51
Tryptophan	23	23	24	23	23

Each batch of neurotoxin was acid hydrolyzed for 24, 48 and 72 hr.
*1/2 Cys in batches 1 and 2 were based on the value from batch 3 (run in duplicate).
Maximum deviation from the mean:
   < 3%   Asp, Pro, ½ Cys, Met, Ile, Leu, Phe, Lys, Arg;
   > 3% but < 5% Thr, Ser, Glu, Ala, Tyr, Trp;
   > 5%   Gly, His, Val.

yielded significantly high or low values (i.e. beyond the instrument error), very good agreement was found between the other two batches (the last column in Table 2). There is no consistent difference among the three batches that can be ascribed to the source of the stock bacterial culture or method of purification.

The type F botulinum neurotoxin independently purified in two laboratories (YANG and SUGIYAMA, 1975; OHISHI and SAKAGUCHI, 1975) have not previously been compared. We found that the neurotoxin purified by these two methods is (i) comparable in purity (as judged by PAGE-SDS) and (ii) similar in amino acid composition. This observation, seemingly trivial, is indeed significant, because the literature on botulinum neurotoxin is replete with claims of purification and characterization of types A, B and E neurotoxin, many of which are inconsistent and irreproducible (see discussions in DASGUPTA and SUGIYAMA, 1977a; DASGUPTA and RASMUSSEN, 1983).

REFERENCES

DASGUPTA, B. R. (1981) Structure and structure function relation of botulinum neurotoxins. In: Biomedical Aspects of Botulism, p. 1 (LEWIS, G. E. JR., Ed.). New York: Academic Press.
DASGUPTA, B. R. and RASMUSSEN, S. (1983) Purification and amino acid composition of type E botulinum neurotoxin. Toxicon 21, 535.
DASGUPTA, B. R. and SUGIYAMA, H. (1977a) Biochemistry and pharmacology f botulinum and tetanus neurotoxins. In: Perspectives in Toxinology, p. 87 (BERNHEIMER, A. W., Ed.). New York: John Wiley.

DasGupta, B. R. and Sugiyama, H. (1977b) Single chain and dichain forms of neurotoxin in type F *Clostridium botulinum* culture. *Toxicon* 15, 466.

Edelhoch, H. (1967) Spectroscopic determination of tryptophan and tyrosine in pr teins. *Biochemistry* 6, 1948.

Lewis, G. E., Jr., Kulinski, S. S., Reichard, D. W. and Metzger, J. F. (1981) Detection of *Clostridium botulinum* type G toxin by enzyme-linked immunoabsorbent assay. *Appl. environ. Microbiol.* 42, 1018.

Moore, S. (1972) The precision and sensitivity of amino acid analysis. In: *Chemistry and Biology of Peptides, Proceedings of the Third American Peptide Symposium*, p. 629 (Meienhofer, J., Ed.). Ann Arbor, Michigan: Ann Arbor Publishers.

Moore, S. and Stein, W. H. (1963) Chromatographic determination of amino acids by the use of automatic recording equipment. *Meth. Enzym.* 6, 819.

Ohishi, I. and Sakaguchi, G. (1975) Molecular construction of *Clostridium botulinum* type F progenitor toxin. *Appl. Microbiol.* 29 444.

Sugiyama, H. (1980) *Clostridium botulinum* neurotoxin. *Microbiol. Rev.* 44, 419.

Syuto, B. and Kubo, S. (1981) Separation and characterization of Heavy and Light chains from *Clostridium botulinum* Type C toxin and their reconstitution. *J. biol. Chem.* 256, 3712.

Weber, K. and Osborn, M. (1969) The reliability of molecular weight determination by dodecyl sulfate-polyacrylamide gel electrophoresis. *J. biol. Chem.* 244, 4406.

Yang, K. H. and Sugiyama, H. (1975) Purification and properties of *Clostridium botulinum* Type F toxin. *Appl. Microbiol.* 29, 598.

Please provide the following references

```
L1    ANSWER 13 OF 17      MEDLINE
AN    97016817      MEDLINE
DN    97016817    PubMed ID: 8863443
TI    Organization and phylogenetic interrelationships of genes encoding
      components of the botulinum toxin complex in proteolytic Clostridium
      botulinum types A, B, and F: evidence of chimeric sequences in the gene
      encoding the nontoxic nonhemagglutinin component.
AU    East A K; Bhandari M; Stacey J M; Campbell K D; Collins M D
CS    Department of Microbiology, Institute of Food Research, Reading,
      Berkshire, United Kingdom.. alison.east@bbsrc.ac.uk
SO    INTERNATIONAL JOURNAL OF SYSTEMATIC BACTERIOLOGY, (1996 Oct) 46 (4)
      1105-12.
      Journal code: 0042143. ISSN: 0020-7713.
CY    United States
DT    Journal; Article; (JOURNAL ARTICLE)
LA    English
FS    Priority Journals
OS    GENBANK-X78230; GENBANK-X87974; GENBANK-X92973; GENBANK-X96491;
      GENBANK-X96492; GENBANK-X96493; GENBANK-X96494
EM    199611
ED    Entered STN: 19961219
      Last Updated on STN: 19980206
      Entered Medline: 19961115
```

```
L1    ANSWER 14 OF 17      MEDLINE
AN    94297488      MEDLINE
DN    94297488    PubMed ID: 7764998
TI    Conserved structure of genes encoding components of botulinum
      neurotoxin complex M and the sequence of the gene coding for the
      nontoxic component in nonproteolytic Clostridium botulinum type
      F.
AU    East A K; Collins M D
CS    Department of Microbiology, Institute of Food Research, Reading
      Laboratory, UK.
SO    CURRENT MICROBIOLOGY, (1994 Aug) 29 (2) 69-77.
      Journal code: 7808448. ISSN: 0343-8651.
CY    United States
DT    Journal; Article; (JOURNAL ARTICLE)
LA    English
FS    Biotechnology
EM    199408
ED    Entered STN: 19950809
      Last Updated on STN: 19970203
      Entered Medline: 19940810
```

```
L1    ANSWER 15 OF 17      MEDLINE
AN    93012902      MEDLINE
DN    93012902    PubMed ID: 1398040
TI    Sequence of the gene encoding type F neurotoxin of
      Clostridium botulinum.
AU    East A K; Richardson P T; Allaway D; Collins M D; Roberts T A; Thompson D
      E
```

1

Organization and Phylogenetic Interrelationships of Genes Encoding Components of the Botulinum Toxin Complex in Proteolytic *Clostridium botulinum* Types A, B, and F: Evidence of Chimeric Sequences in the Gene Encoding the Nontoxic Nonhemagglutinin Component

ALISON K. EAST,* MANJU BHANDARI, JOANNA M. STACEY,
KATHRYN D. CAMPBELL, AND MATTHEW D. COLLINS

*Department of Microbiology, Institute of Food Research, Reading,
Berkshire RG6 6BZ, United Kingdom*

The cluster of genes encoding components of the botulinum neurotoxin (BoNT) complex was mapped in proteolytic (group I) *Clostridium botulinum* strains encoding BoNT types A, B, and F. Two different arrangements of genes were found: type A strain 62A and type B strain NCTC 7273 have similar organizations of genes encoding BoNT, the nontoxic nonhemagglutinin component (NTNH), hemagglutinin components, and P-21; type F strain Langeland has genes encoding BoNT, NTNH, and P-21, and a previously unidentified open reading frame encoding a protein of 416 amino acids. A group of type A strains typified by infant strain Kyoto-F, which is unlike type A strain 62A, lacks genes for hemagglutinin components and exhibits an organization similar to that of type F. Sequencing and pairwise analysis revealed the presence of possible chimeric sequences in some NTNH genes of proteolytic *C. botulinum*. Discordance in genealogical trees derived from different regions of the NTNH genes was observed which could be symptomatic of recombination and which may indicate that the NTNH gene represents a hot spot for such events within the cluster of genes encoding the BoNT complex. It is also evident that the phylogenetics of the NTNH gene, which is linked to the gene encoding BoNT, does not mirror the evolutionary history of the BoNT, upon which the *C. botulinum* species complex is defined and subdivided.

Strains of the species *Clostridium botulinum* produce the extremely toxic botulinum neurotoxin (BoNT). BoNT is the causative agent of the severe neuroparalytic illness in humans and animals referred to as botulism. BoNTs are classified into seven different types (A through to G) by reaction with specific antisera. *C. botulinum* strains are traditionally divided into four physiological groups, which are designated I to IV (18). Proteolytic group I *C. botulinum* strains produce a single toxin or, occasionally, dual toxins of type A, B, or F. Nonproteolytic group II *C. botulinum* strains invariably produce a single toxin of type B, E, or F. Group III strains produce BoNT type C or BoNT/D and are responsible for animal botulism, while group IV strains produce BoNT/G. Strains of groups I and II, especially those of types A and B, are important agents of both food-borne and infant botulism (18). In addition, some strains of the non-*C. botulinum* species are known to synthesize BoNT (viz., *Clostridium barati* [type F] and *Clostridium butyricum* [type E]). Recently, the evolutionary interrelationships of *C. botulinum* types A to G, *C. barati* type F, and *C. butyricum* type E and their BoNT genes have been systematically investigated. 16S rRNA sequencing revealed the presence of four distinct phylogenetic lines within *C. botulinum* which correlate with the four physiological groups (I to IV) (22). The same studies revealed that neurotoxigenic *C. barati* and *C. butyricum* strains are authentic members of their respective species and are genealogically distinct from each other and the four *C. botuli-*

* Corresponding author. Mailing address: Department of Microbiology, Institute of Food Research, Earley Gate, Whiteknights Road, Reading, Berkshire RG6 6BZ, United Kingdom. Phone: 44 1734 357225. Fax: 44 1734 267917. Electronic mail address: alison.east @bbsrc.ac.uk.

num lines (22). However, major differences in the pattern of relationships inferred from 16S rRNA and BoNT gene sequences were observed that were indicative of BoNT gene transfer between *C. botulinum* rRNA lines and the non-*C. botulinum* species (3, 10, 20).

BoNT is produced by *C. botulinum* as a noncovalently bound complex of two or more protein components. The smallest complex, the M complex (ca. 300 kDa), which is found in all strains (except type G) which produce neurotoxin, consists of BoNT with a component that has a similar size (ca. 150 kDa), the nontoxic nonhemagglutinin component (NTNH) (33). The size of the complex formed by *C. botulinum* shows some correlation with the toxin type encoded; strains encoding types E and F (26) and some type A strains (30) produce only the M complex, while strains encoding types A, B, C, and D, in addition, produce L complexes (ca. 500 kDa), which have associated hemagglutinin (HA) activity (28). Type G strains produce only the L complex (25). Many type A strains produce not only the M and L complexes, but also a larger LL complex (ca. 900 kDa). The purported function of the nontoxic components of the complex is to increase the stability of BoNT in the gastrointestinal tract (29). The gene encoding NTNH is located immediately upstream of that of BoNT in all toxin types (6), and genes for other components of the complexes have been shown clustered upstream of the NTNH gene in strains encoding types A and B (8) and type C (15, 19, 32, 33).

Recently, we showed that two distinct populations of *C. botulinum* type A strains exist, encoding distinct BoNT/A genes (designated BoNT/A1 and BoNT/A2) which can be distinguished from each other by DNA probes and by restriction fragment length polymorphism analysis (4). Type A1 and A2 strains encode BoNTs which show ca. 89% sequence identity

(1, 31, 35). These strains also differ in the composition of the progenitor toxin complex produced and the organization of genes in the toxin cluster (4). It has been shown that approximately 50% of *C. botulinum* type A strains contain unexpressed BoNT/B gene sequences (12). In the *C. botulinum* strains thus far examined, those which carry additional BoNT/B genes are invariably of type A1 (4). To date, knowledge of the gene organization and evolutionary interrelationships of the various progenitor toxin components has not been systematically examined. Since genes encoding these complex components are physically adjacent to genes responsible for BoNT production, it is not unreasonable to expect that their genealogical relationships would mirror that of the neurotoxin.

In this paper, we report elucidation of the organization of the cluster of genes encoding components of the BoNT complex in group I *C. botulinum* type A, B, and F strains. We report the finding of two distinct gene organizations: one for types A1 and B and another for types A2 and F. The complete sequence of the NTNH gene is presented for strains encoding BoNT types A1, A2, and B, and genealogical analyses provide evidence for chimera-like sequences in some genes. Discordance in trees derived from different regions of the NTNH genes, which is symptomatic of recombination, was observed, indicating that the NTNH gene may be a hot spot for such events within the cluster of genes coding for the BoNT complex.

MATERIALS AND METHODS

Bacterial strains and plasmids. Genomic DNA was prepared from the clostridial strains listed in Table 1 according to the method described by Lawson et al. (23). *Escherichia coli* One Shot obtained from Invitrogen (R&D Systems, Cowley, United Kingdom) was used as the recipient for cloning.

PCR and cloning. PCR was carried out with Perkin-Elmer Ampli-*Taq* DNA polymerase (Applied Biosystems, Warrington, United Kingdom) as described previously (6) according to the scheme outlined in Fig. 1. The sequences of primers used are given in Table 2. PCR products were cloned with a TA cloning kit (Invitrogen) as described previously (6).

Hybridization analysis. Hybridizations were carried out with an ECL kit (Amersham International, Amersham, United Kingdom) according to the manufacturer's instructions. Slot blots were performed essentially as described by Campbell et al. (2), with hybridization carried out at 42°C and washing carried out at 50°C. The oligonucleotide hybridization probes used are listed in Table 2.

DNA sequencing. Double-stranded DNA was sequenced as described previously (2) with a U.S. Biochemicals Sequenase kit (Amersham) according to the manufacturer's instructions. Clones obtained from two separate PCRs were sequenced, and, in the case of discrepancy, a third independent clone was analyzed.

Analysis of data. Analysis was performed with the molecular biology software of University of Wisconsin Genetics Computer Group package (5). The neighbor-joining method described by Saitou and Nei (27) was used for construction of the phylogenetic trees. The stability of trees was assessed by bootstrap analysis (11). Bootstrap values were calculated from 250 replicates.

Nucleotide sequence accession numbers. The sequences presented here have been given EMBL/GenBank accession numbers as follows: NTNH gene sequences from strains 62A, Kyoto-F, and NCTC 7273, respectively, X92973, X87974, and X78230; P-21-NTNH gene sequences from strains Kyoto-F and Langeland, respectively, X96493 and X96494; P-21 partial gene sequences from strains Chiba-H and NCTC 9837, respectively, X96491 and X96492.

RESULTS

Analysis of the NTNH genes. The regions upstream of the BoNT genes in type A strains 62A and Kyoto-F, type B strain NCTC 7273, and type F strain Langeland were amplified by PCR, and fragments were cloned as outlined in Fig. 1. The gene encoding NTNH was found immediately upstream of BoNT, and its sequences in types A and B were determined. In type A strain 62A, the C-terminal 103 amino acids correspond to translation of the region upstream of BoNT/A reported by Binz et al. (1), and for type B strain NCTC 7273, the C-terminal 8 amino acids correspond to translation of the region upstream of BoNT/B (34). For type A strain Kyoto-F, the

TABLE 1. Details of strains used and results of PCR and probing experiments

Strain[a]	PCR with H1-H2	Probe HA33/A	PCR with P1-P2	PCR with I+-NN2	PCR with I--NN2
C. botulinum group I					
NCTC 887 (A1)	+	+	+	+	−
599 (A1)	+	+	+	+	−
603 (A1)	+	+	+	+	−
2775 (A1)	+	+	+	+	−
62A (A1)	ND[b]		+	+	−
547 (A1)	+	+	+	+	−
726 (A1)	+	+	+	+	−
1690 (A1)	+	+	+	ND	ND
NCTC 7272 (A1)	ND		+	ND	ND
NCTC 2916 [A1(B)]	+	+	+	+	+
NCTC 11199 [A1(B)]	+	+	+	+	+
2119 [A1(B)]	+	+	+	+	+
5311 [A1(B)]	+	+	+	+	+
7826 [A1(B)]	+	+	+	+	+
MDa10 [A1(B)]	+	+	+	+	+
588 [A1(B)]	+	+	+	+	+
13280 [A1(B)]	+	+	+	+	+
667 [A1(B)]	+	+	+	+	+
657 (A1/B)	+	+	ND	ND	ND
3281(Bf)	+	+	ND	+	+
Kyoto-F (A2)	ND		+	−	+
Chiba-H (A2)	−		+	−	+
NCTC 9837 (A2)	−		+	−	+
Y-8036 (A2)	−		−	−	+
7105-H (A2)	−		−	−	+
7103-H (A2)	−		−	−	+
KZ1828 (A2)	−		−	−	+
NCTC 2012 (A2)	−		−	−	+
Langeland (F)	−		+	−	+
C. botulinum group II					
Eklund 2B (B)	+	−	+	+	−
Scott 2129B (B)	+	−	+	+	−
Colworth 151 (B)	+	−	+	+	−
Hobbs FT50 (B)	+	−	+	+	−
Hazen 36208 (E)	−		−	−	+
Eklund 202F (F)	−		−	−	+
VPI 2093-1 (NT)	−		−	−	−
C. botulinum group IV					
ATCC 27322 (G)	−		+	−	−
C. barati					
ATCC 43756 (F)	−		−	−	−
C. butyricum					
ATCC 43755 (E)	−		−	−	+

[a] Toxin types are given in parentheses after strain designations. A1(B), production of BoNT/A but with silent BoNT/B present on the genome; A1/B, BoNT/A and BoNT/B (dual toxin) production; Bf, BoNT/B and BoNT/F production, but with BoNT/F produced only in minor amounts; NT, nontoxic.
[b] ND, not determined.

C-terminal 16 amino acids of NTNH correspond to the sequence upstream of BoNT/A reported by Willems et al. (35). The proteins encoded by strain 62A (type A1), Kyoto-F (type A2), and NCTC 7273 (type B) sequences consisted of 1,193, 1,159 and 1,197 amino acids, respectively, and showed homology to published NTNH sequences of toxins of other types (6, 13, 14, 33). While this work was in progress, Fujita et al. (16) reported the NTNH sequence of type A strain A-NIH, which is identical to that reported here for type A strain 62A. The overall homology shown between these NTNHs (>60% iden-

(i) Type A

A1: strain 62A

A2: strain Kyoto-F

(ii) Type B

strain NCTC 7273

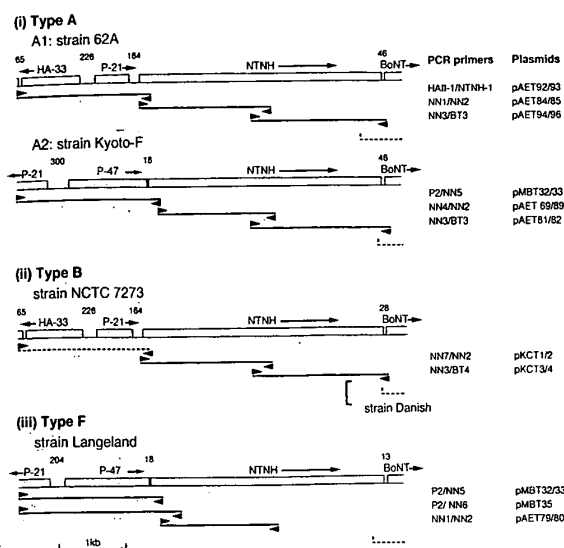strain Danish

(iii) Type F

strain Langeland

FIG. 1. Strategy for PCR amplification and cloning of fragments encoding components of the BoNT complex from proteolytic *C. botulinum* type A, B, and F strains. The strains used and the approximate positions of binding of PCR primers are shown. Primer sequences are given in Table 2. The positions of genes encoding the components of the complex and direction of transcription are shown. The sizes of the intergenic regions are indicated. Dotted lines, sequences previously reported (type A strains 62A [1] and Kyoto-F [35], type B strains NCTC 7273 [8] and Danish [34], and type F strain Langeland [10]).

evident. The largest and most striking of these is at position 115, where there are an extra 33 amino acids in types A1 and B compared with A2. The extra 33 residues are present in the NTNH sequence of *C. botulinum* type C (33) but are absent from those of *C. botulinum* type E (14), *C. butyricum* (13), and nonproteolytic *C. botulinum* type F (6). Preliminary sequence analysis of the region in proteolytic *C. botulinum* type F strain Langeland shows that it too lacks these 33 amino acids (9). Since all of the strains examined which possess this insert (types A1, B, and C) have HA components in their progenitor toxin complex, it is interesting to speculate that this part of the NTNH may somehow be involved in interaction with HAs. It has recently been shown that the NTNH component of type A, C, and D strains of the M complex has been proteolytically cleaved at a position within the extra 33 amino acids, while the NTNH component of the L complex is unprocessed (16). Since the L complex contains components which have HA activity, while the M complex lacks them, processing within this region may prevent (directly or indirectly) the HA components from interacting to form the larger complex.

To examine the occurrence of this insert in the NTNH genes of other strains, PCR primer I+, which binds to a region within the extra 33 residues (positions 122 to 127 [Table 2]), and primer I−, which spans the site at which the additional residues are present in some strains (residues 110 to 117 for NTNH/A2 [Table 2]), were used in combination with conserved primer NN2 (positions 692 to 699 [Table 2]). PCR products were obtained with primer combination I+-NN2 with all strains encoding type A1 [including A1(B)] and B tested but not with any of those encoding type A2 (Table 1). With primer combination I−-NN2, positive results were obtained for all strains encoding types A2, E, F, and, unexpectedly, all type A strains containing silent BoNT/B sequences [i.e., A1(B)]. Since products were obtained with both pairs of primers for type A1(B) strains, this indicates that NTNH genes encoding proteins with and without the additional 33 amino acids are

tity) is higher than that shown between the BoNTs, which can be as low as ca. 34% identity (3, 10). From an alignment of the NTNH amino acid sequences (data not shown), it is clear that although the proteins show a high degree of homology, several insertions and/or deletions of one or more amino acids are

TABLE 2. Oligonucleotide primers used in PCR and probing experiments

PCR primer or probe	Sequence 5'-3'	Target amino acids[a]	Target positions[b]
PCR primers			
BT3	TGCTTCTGGTGGTGGATTTAAATCTCC	GDLNPPPEA	57–65 (BoNT/A)
BT4	CTCCCCGTACCTCTCGCAAATGG	PFARGTG	25–31 (BoNT/B)
H1	AAATTGTTACCATCTCCTGTAAGGC	IVTISCKA	16–23 (HA-33)
H2	AATTAAATACTTGAATAGCAGTTCCGT	GTAIQVFN	275–282 (HA-33)
HAII-1	AT(A/T)TATAATTACCATTAGG	PNGYNI	9–15 (HA-17)
I+	ATATGATTACTTTTGGATCAGCAC	MITFGSA	122–127 (NTNH/A1)
I−	CTTATGAATATAAACCTGGAGATTAT	YEYKPGDY	110–117 (NTNH/A2)
NN1	ATA(A/G)ATTC(A/T)CCAGTGGA(T/C)AA	I(N/D)SPVDN	9–15 (NTNH)
NN2	ATTTTCTTGTGCTAATATTGA	SILAQEND	692–699 (NTNH)
NN3	TTAGTTTCTTAGATCAATGGTGG	SFLDQWW	669–675 (NTNH)
NN4	AGAAAGGTATTATGGAGACA	ERYYGE	44–49 (NTNH)
NN5	TTAGAATCATAAATTCC(T/A)CC(G/A)TC	DGGIYDS	60–66 (NTNH)
NN6	TGATTATCTTTAGATACT	VSKDN	157–161 (NTNH)
NN7	TAAGGGGGCGTGTATAGTA	Noncoding	Upstream of NTNH
NTNH-1	CTTGCTCTAACTATTACAAC	VVIVRAR	18–24 (NTNH)
P1	GATAT(A/T)(C/G/T)TG(A/T)(A/G)(T/C)CAT(C/T)T(A/T)TGG	DI(L/V)xH(F/L)W	44–50 (P-21)
P2	T(T/C)TTATATA(C/T)AG(A/C)(T/C)TGACGACT	SRQ(S/A)(V/I)YK	167–173 (P-21)
Probes			
H33/A	CTTCTAATGATCAAAATA	SNDQN	231–235 (HA-33)
H33/Bp	TCCCGGTAACGGTAACGT	PGNGNV	33–38 (HA-33)
A33/Bnp	GCTATGACAGATCTAAG	AMTDLS	174–179 (HA-33)

[a] Amino acids for oligonucleotide target positions. x, any amino acid.
[b] Binding positions are given with reference to the following: BoNT/A (1), BoNT/B (34), HA-33 (8), HA-17 (19), NTNH (6), and P-21 (Fig. 6).
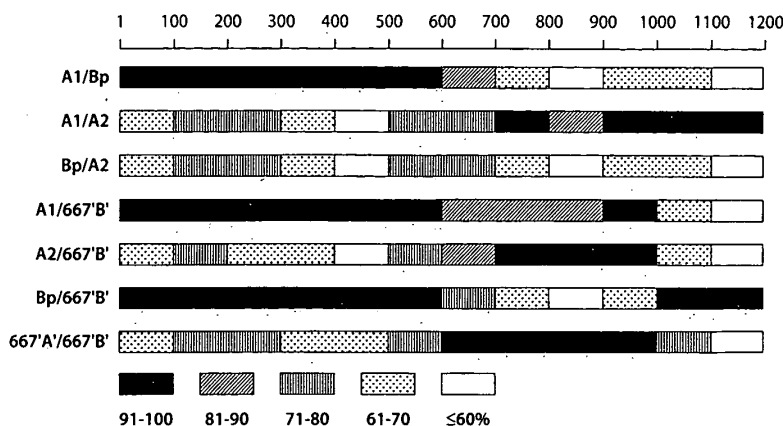
FIG. 2. Pairwise comparison of NTNH sequences from proteolytic *C. botulinum* type A strains 62A (A1) and Kyoto-F (A2), type B strain NCTC 7273 (Bp), and type A(B) strain 667. 667'A', A cluster; 667'B', B cluster (21). Comparisons were performed on blocks of 100 amino acids (shown at the top). Shading indicates the degrees of homology.

present. Hutson et al. (21) have sequenced the two BoNT gene clusters encoded by *C. botulinum* type A1(B) strain 667, which revealed that the NTNH gene of the A cluster (i.e., upstream of the expressed BoNT/A) lacks the extra 33 amino acids while the NTNH encoded by the B cluster possesses them.

Pairwise analyses were performed on the NTNH gene sequences of proteolytic *C. botulinum*. It was evident from these analyses that for some pairs of sequences there was considerable variation in the degree of homology along their lengths. The variability in the pattern of relationship between different positions in the NTNH sequences of representative combinations of *C. botulinum* strains is illustrated in Fig. 2. For example, the NTNHs of proteolytic *C. botulinum* types A and B show very high sequence relatedness for approximately the first 600 residues (>99% identity). In contrast, the amino acids from position ca. 600 to the C terminus have a much-reduced relatedness, ca. 66% identity. The reverse is observed for the NTNHs of *C. botulinum* types A1 and A2, with the C-terminal end showing high sequence identity (ca. 93% for position 670 to the end) compared with that for the N-terminal region (ca. 67% for positions 1 to 669). Matrices of pairwise sequence homology data from different regions of the NTNH (residues 1 to 550, 551 to 1020, and 1021 to the C terminus) were used to construct comparative genealogical trees (Fig. 3). Major differences in the three trees reinforce the incongruence in the pattern of relationships along the length of the NTNHs. Such marked discordance could be symptomatic of recombination and suggests that the NTNH gene is a hot spot for such events within the BoNT complex cluster of genes. Hutson et al. (21) reported that the NTNH gene adjacent to the silent BoNT/B gene (B cluster) in *C. botulinum* type A(B) strain 667 was chimeric, with the 5' and 3' regions of the gene exhibiting high homology with the corresponding regions of the type B NTNH gene (Fig. 3), while a 417-amino-acid sequence in the central region was identical to that in the NTNH whose gene was located adjacent to the expressed BoNT/A gene (A cluster) (Fig. 2). The results of the present study demonstrate that such chimera-like NTNH genes may be commonplace among proteolytic *C. botulinum*.

A comparison of the homology values along the length of the cluster of genes encoding the BoNT complex for type A strains 62A and Kyoto-F and type B strain NCTC 7273 is shown in Fig. 4. In addition to the variation in the degree of homology in the coding regions, further weight is given to the idea of

recombination by observation of the intergenic regions. In these strains, the intergenic regions are of identical size and (almost) identical sequence between HA-17 and HA-33, HA-33 and P-21, and P-21 and NTNH in type A strain 62A and type B strain NCTC 7273, while the corresponding region in type A strain Kyoto-F is noncomparable (Fig. 4). In contrast, the intergenic region between the NTNH and BoNT genes is identical in size and sequence in the two type A strains, while that of type B strain NCTC 7273 is different (Fig. 4).

Investigation of the HA-33 genes of types A and B. We recently showed that type A strain NCTC 7272 and type B strains NCTC 7273 and 17B encode genes for HA-17, HA-33, and P-21 (linked in this order) upstream of those for NTNH and BoNT (Fig. 1) (8). The region encoding HA-33 and P-21 was cloned from type A strain 62A (Fig. 1) and was found to be identical to that from type A strain NCTC 7272 (8). To determine if the HA-33 gene is present in other strains, PCR primers based on conserved amino acids were developed (primers H1 and H2 [Table 2]). Positive results (PCR product of ca. 1.2
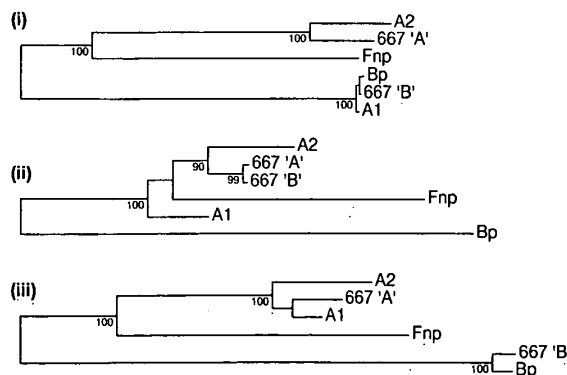
FIG. 3. Trees revealing the discordance in relationships along the lengths of NTNHs, residues 1 to 550 (i), residues 551 to 1020 (ii), and residue 1021 to C terminus (iii). NTNH sequences from the following were compared: proteolytic *C. botulinum* type A strains 62A (A1) and Kyoto-F (A2), type B strain NCTC 7273 (Bp), type A(B) strain 667, A cluster (667'A') and B cluster (667'B') (21), and nonproteolytic *C. botulinum* type F strain Eklund 202F (Fnp) (6). The numbers on the tree indicate bootstrap values for branch points; only values of >90 are shown.
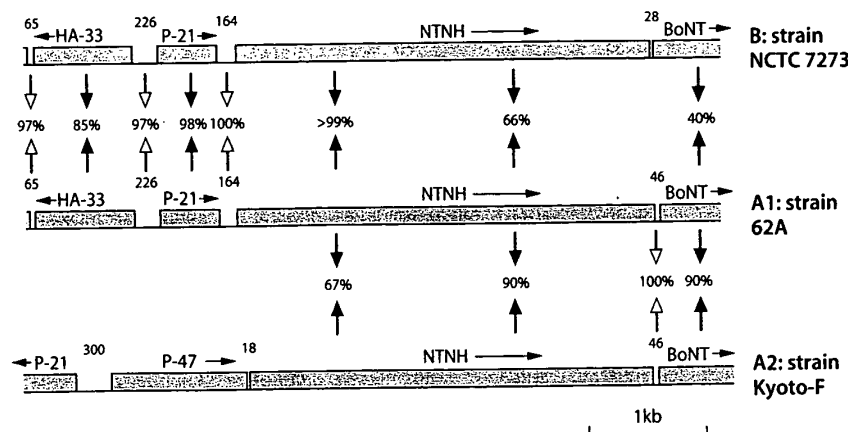
FIG. 4. Homology along the length of the BoNT gene cluster for *C. botulinum* type A and B strains. Closed arrows, amino acid identity values for proteins encoded by the genes, open arrows, nucleic acid identity values for intergenic regions. Values for NTNH refer to residues 1 to 600 and 601 to the C terminus.

kb) were obtained for *C. botulinum* strains encoding type A1 [including A1(B)] and B but not for those encoding type A2, E, F, and G or toxigenic strains of *C. barati* and *C. butyricum* (Table 1). Although the lack of a PCR product may be due to the mismatch of primers with the target gene, these results correlate with the presence or absence of HA activity in the progenitor complexes; strains encoding types A and B, which produce complexes with HA activity, give positive results, while strains which lack HA activity (types A2, E, and F and strains of *C. barati* and *C. butyricum*) give no PCR product. Although type G strains produce complexes which contain components with HA activity, purification of the components shows that none correspond to HA-33 of type A, B, or C (25).

Hybridization of the PCR products with probe H33/A (Fig. 5A) shows a positive result for all type A1, A1(B), A1/B, and Bf strains. In contrast, the probe designed to detect HA-33 of proteolytic B (H33/Bp) gave a signal only with the positive control and not with any other strains, including the type A1(B) strains (Fig. 5C). The probe to detect HA-33 of nonproteolytic type B (H33/Bnp) hybridized only with the PCR product from the nonproteolytic B strains (Fig. 5B). Interestingly, all of the A1(B) strains examined hybridized only with probe H33/A, indicating the presence of only type A HA-33. Sequence analysis of *C. botulinum* type A1(B), strain 667, showed only one HA-33 gene to be present; although it was located in the B cluster, it showed the highest homology with
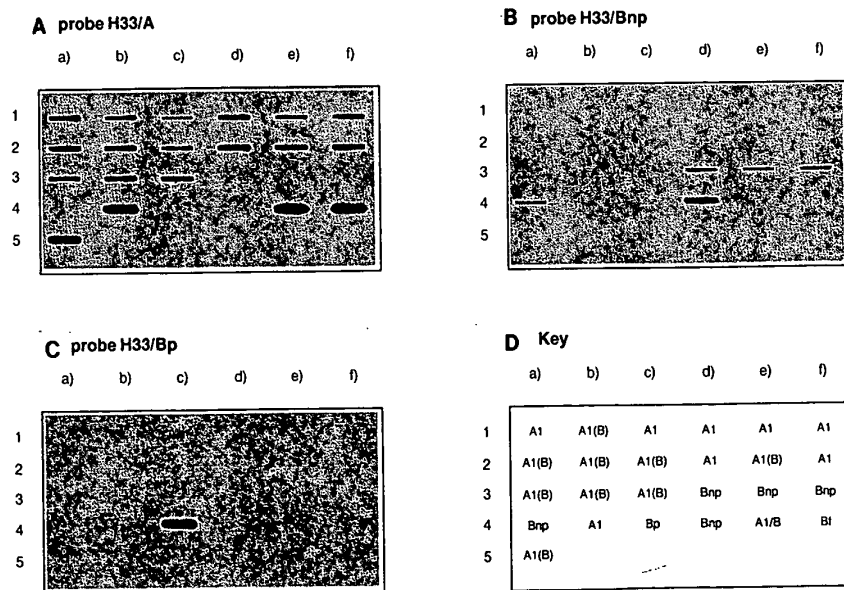
FIG. 5. (A to C) Slot blots with probes designed to detect the genes encoding HA-33 type A, HA-33 type B from nonproteolytic *C. botulinum*, and HA-33 type B from proteolytic *C. botulinum*, respectively; (D) key with strain type, showing hybridization with the following strains: 547 (1a), 588 (1b), 599 (1c), 603 (1d), 726 (1e), NCTC 887 (1f), 5311 (2a), 2119 (2b), NCTC 2916 (2c), 2775 (2d), 7826 (2e), 1690 (2f), NCTCC 11199 (3a), 13280 (3b), MDa10 (3c), Scott 2129B (3d), Colworth 151 (3e), Hobbs FT50 (3f), Eklund 2B (4a), pAET70 (cloned HA-33 gene from type A strain NCTC 7272) (8) (4b), pAET59 (cloned HA-33 gene from proteolytic type B strain NCTC 7273) (8) (4c), pAET55 (cloned HA-33 gene from nonproteolytic type B strain Eklund 17B) (8) (4d), 657 (4e), 3281 (4f), and 667 (5a).
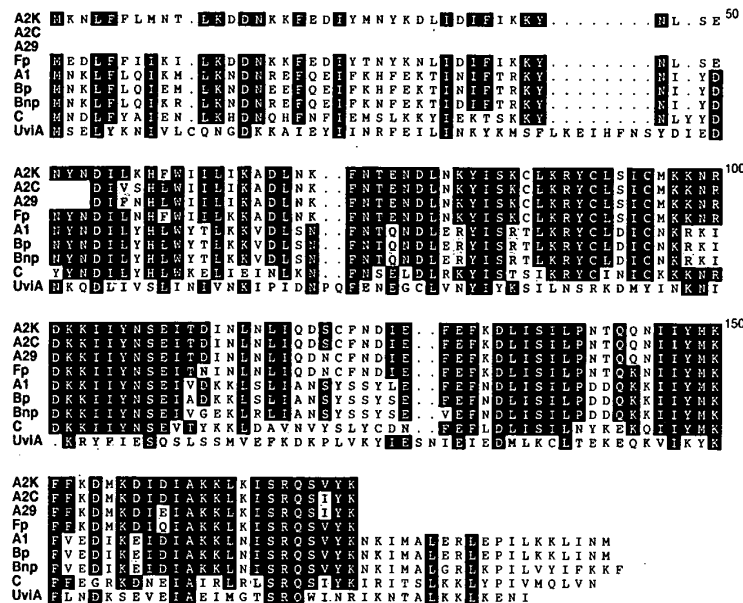
FIG. 6. Comparison of derived amino acid sequences for P-21s from *C. botulinum* type A strains 62A (A1), Kyoto-F (A2K), Chiba-H (A2C), and NCTC 9837 (A29); type B strains NCTC 7273 (Bp), and Eklund 17B (Bnp) (8); type C strain 468 (19); type F strain Langeland; and UviA from *C. perfringens* (17). Black and shaded areas, identical and similar amino acids in five or more sequences, respectively; dots, absent residues.

HA-33/A (ca. 99% identity) and possessed the hybridization site of the H33/A probe (21). The single Bf strain examined (strain 3281) hybridized only with the HA-33 type A probe (Fig. 5). Further experiments are under way to investigate this result.

Elucidation of P-21 gene sequences in *C. botulinum* types A2 and F. A gene encoding a potential regulator of the BoNT complex was reported for *C. botulinum* types A and B (designated *P-21*) (8) and type C (designated *orf-22*) (19). This gene is clustered with those of the components of the BoNT complex, although it maps to different positions in strains of different toxin types. PCR primers based on conserved regions of *P-21* (P1 and P2 [Table 2]) were designed and used to investigate the presence of this gene in other strains (Table 1). No PCR products were obtained for any of the strains of nonproteolytic *C. botulinum* type F, *C. botulinum* type E, *C. barati*, or *C. butyricum* examined. PCR amplification products with the expected size (ca. 400 bp) were obtained for all strains encoding types A1, A1(B), B, and G for the single proteolytic *C. botulinum* type F strain and three of the type A2 strains examined (Table 1). The PCR amplification products from the proteolytic *C. botulinum* type F strain (Langeland) and the three type A2 strains (Kyoto-F, Chiba-H, and NCTC 9837) were cloned, and their sequences were determined. Very high sequence homology was observed between the P-21 proteins from these type A2 strains and type F strain Langeland (ca. 91% sequence identity). The P-21 proteins from type A2 strains and type F strain Langeland show ca. 63% amino acid identity with the P-21 proteins from *C. botulinum* type A1 and type B strains (proteolytic and nonproteolytic), ca. 55% sequence identity with the *orf-22* gene product from *C. botulinum* type C (19), and ca. 26% identity with the *uviA* gene product from *Clostridium perfringens* (17). UviA is thought to be a DNA-binding protein with a possible regulatory function (24). Figure 6 shows a comparison of the sequences of the P-21

proteins of *C. botulinum* and UviA from *C. perfringens*. To determine the position of the P-21 gene in strains Langeland and Kyoto-F, PCR strategies were employed. The P-21 gene was found to map in an identical position ca. 2 kb upstream of the NTNH gene in both of these strains. The orientation of P-21 is such that it is transcribed away from the 5' end of the NTNH gene (Fig. 1).

Analysis of a new open reading frame. In an earlier study, we reported the 3' end of an unknown *orf* located upstream of NTNH gene in nonproteolytic *C. botulinum* type F strain Eklund 202F (6). Cloning and sequencing the region between the NTNH and P-21 genes revealed an *orf*, which was preceded by a putative Shine-Dalgarno sequence, encoding a protein of 416 amino acids in both proteolytic *C. botulinum* type F (Langeland) and type A2 (Kyoto-F) strains. This protein has been designated P-47, since it has relative molecular weight values of 47,476 in strain Kyoto-F and of 47,420 in strain Langeland. A comparison of the derived amino acid sequences of P-47 proteins for strains Kyoto-F and Langeland shows ca. 79% identity overall but ca. 96% identity in the first 200 amino acids. The translated 3' end of the *orf* in *C. botulinum* type F strain Eklund 202F shows homology to the C termini of P-47 proteins encoded in strains Langeland and Kyoto-F (ca. 66 to 69% identity over 271 amino acids [7]). Interestingly, the 3' end of this *orf* is also found in *C. botulinum* type A(B) strain 667 (21), and translation of sequence upstream of the NTNH gene in *C. botulinum* type E (14) and *C. butyricum* (13) reveals an *orf* encoding a protein with homology to P-47s of Kyoto-F and Langeland (ca. 54 to 80% identity). This gene maps in identical positions in these strains (Fig. 1). The gene encoding P-47 is, therefore, widely distributed among strains encoding BoNT and is linked or clustered with genes encoding neurotoxin complex components. It may be pertinent that this gene has been found only in toxin gene clusters which lack genes for HA components of the complex.

DISCUSSION

It has long been recognized that strains of *C. botulinum* are physiologically heterogeneous, and four distinct phenotypic groups (I to IV) are recognized (18). These four metabolically distinct groups do not, however, necessarily correlate with the serological specificities of the BoNT produced (18). Genealogical analysis based on 16S rRNA gene sequences has shown that *C. botulinum* consists of four highly divergent lines which correlate with the four phenotypic divisions (22). Recent sequencing studies of the genes encoding BoNT have unequivocally demonstrated the marked discordance between trees depicting natural relationships (i.e., 16S rRNA) of the organisms and those of the various BoNT types produced (e.g., see references 3, 10, and 20). In particular, the results of these comparative sequencing studies are strongly indicative of BoNT gene transfer between *C. botulinum* metabolic groups or rRNA lines (viz., BoNT/B in groups I and II, BoNT/E in group II and toxigenic *C. butyricum*, and BoNT/F in groups I and II and toxigenic *C. barati*). As outlined earlier, the gene encoding BoNT is clustered with several other genes encoding components of the progenitor complex. Since BoNT is responsible for botulism and BoNT is the primary trait for defining and subdividing the so-called species *C. botulinum*, a thorough understanding of the evolutionary interrelationships of the genes encoding this toxin and other components of the progenitor complex is essential. Knowledge of the organization of the cluster of genes encoding BoNT complex components is very limited. Information on the genealogy of the various progenitor proteins is also unknown, and since genes encoding these are proximal to that responsible for BoNT production, it is likely that their genealogical patterns would be similar to that of the neurotoxins. In this study, we have examined the arrangements of genes encoding components of the BoNT complex in proteolytic (group I) *C. botulinum* type A, B, and F strains. Two different arrangements were found. Type A strain 62A and type B strain NCTC 7273 possessed similar organizations, viz., *BoNT*, *NTNH*, *P-21*, and *HA* components; type F strain Langeland and infant type A strain Kyoto-F exhibited gene organizations similar to each other, viz., *BoNT*, *NTNH*, *P-47*, and *P-21*, and appeared to lack HA genes. In addition to this discordance in gene organization versus BoNT type, we have shown that the genealogy of the NTNH (the gene encoding NTNH is linked to the BoNT gene) does not mirror the evolutionary history of BoNT. Furthermore, we have found evidence of chimeric NTNH gene sequences, which could be symptomatic of recombination, and indicate that the NTNH gene may be a hot spot for such events within the BoNT complex gene cluster. There is clearly a strong case to extend the described studies to embrace further group I strains and *C. botulinum* groups (including other BoNT-producing species) to obtain a better understanding of population structure within this important pathogen and a complete picture of the evolutionary histories of its neurotoxin and associated proteins. Such genetic information is of great relevance to the taxonomy of *C. botulinum*, which is currently defined and subdivided on the basis of the type of neurotoxin produced.

ACKNOWLEDGMENTS

REFERENCES

1. Binz, T., H. Kurazono, M. Wille, J. Frevert, K. Wernars, and H. Niemann. 1990. The complete sequence of botulinum neurotoxin type A and comparison with other clostridial neurotoxins. J. Biol. Chem. **265**:9153–9158.
2. Campbell, K. D., M. D. Collins, and A. K. East. 1993. Gene probes for identification of the botulinal neurotoxin gene and specific identification of neurotoxin types B, E, and F. J. Clin. Microbiol. **31**:2255–2262.
3. Campbell, K. D., M. D. Collins, and A. K. East. 1993. Nucleotide sequence of the gene coding for *Clostridium botulinum* (*Clostridium argentinense*) type G neurotoxin: genealogical comparison with other clostridial neurotoxins. Biochem. Biophys. Acta **1216**:487–491.
4. Cordoba, J. J., M. D. Collins, and A. K. East. 1995. Studies on the genes encoding botulinum neurotoxin type A of *Clostridium botulinum* from a variety of sources. Syst. Appl. Microbiol. **18**:13–22.
5. Devereux, J., P. Haeberli, and D. Smithies. 1984. A comprehensive set of sequence analysis programs for the VAX. Nucleic Acids Res. **12**:387–395.
6. East, A. K., and M. D. Collins. 1994. Conserved structure of genes encoding components of botulinum neurotoxin complex M and the sequence of the gene coding for the nontoxic component in non-proteolytic *Clostridium botulinum* type F. Curr. Microbiol. **29**:69–77.
7. East, A. K., and M. D. Collins. Unpublished data.
8. East, A. K., J. M. Stacey, and M. D. Collins. 1994. Cloning and sequencing of a hemagglutinin component of the botulinum neurotoxin complex encoded by *Clostridium botulinum* types A and B. Syst. Appl. Microbiol. **17**:306–312.
9. East, A. K., J. M. Stacey, S. Hielm, and M. D. Collins. Unpublished data.
10. Elmore, M. J., R. A. Hutson, M. D. Collins, N. J. Bodsworth, S. M. Whelan, and N. P. Minton. 1995. Nucleotide sequence of the gene coding for proteolytic (group I) *Clostridium botulinum* type F neurotoxin: genealogical comparison with other clostridial neurotoxins. Syst. Appl. Microbiol. **18**:23–31.
11. Felenstein, J. 1989. PHYLIP-phylogeny inference package (version 3.2). Cladistics **5**:164–166.
12. Franciosa, G., J. L. Ferreira, and C. L. Hatheway. 1994. Detection of type A, B and E botulism neurotoxin genes in *Clostridium botulinum* and other *Clostridium* species by PCR: evidence of unexpressed type B toxin genes in type A toxigenic organisms. J. Clin. Microbiol. **32**:1911–1917.
13. Fujii, N., K. Kimura, N. Yokosawa, K. Oguma, T. Yashiki, K. Takeshi, T. Ohyama, E. Isogai, and H. Isogai. 1993. Similarity in nucleotide sequence of the gene encoding nontoxic component of botulinum toxin produced by toxigenic *Clostridium butyricum* strain BL6340 and *Clostridium botulinum* type E strain Mashike. Microbiol. Immunol. **37**:395–398.
14. Fujii, N., K. Kimura, N. Yokosawa, T. Yashiki, K. Takeshi, K. Tsuzuki, and K. Oguma. 1993. The complete nucleotide sequence of the gene encoding the nontoxic component of *Clostridium botulinum* type E progenitor toxin. J. Gen. Microbiol. **139**:79–86.
15. Fujinaga, Y., K. Inoue, S. Shimazaki, K. Tomochika, K. Tsuzuki, N. Fujii, T. Watanabe, T. Ohyama, K. Takeshi, K. Inoue, and K. Oguma. 1994. Molecular construction of *Clostridium botulinum* type C progenitor toxin and its gene organization. Biochem. Biophys. Res. Commun. **205**:1291–1298.
16. Fujita, R., Y. Fujinaga, K. Inoue, H. Nakajima, H. Kumon, and K. Oguma. 1995. Molecular characterization of two forms of nontoxic-nonhemagglutinin components of *Clostridium botulinum* type A progenitor toxins. FEBS Lett. **376**:41–44.
17. Garnier, T., and S. T. Cole. 1988. Studies of UV-inducible promoters from *Clostridium perfringens in vivo* and *in vitro*. Mol. Microbiol. **2**:607–614.
18. Hatheway, C. L. 1992. *Clostridium botulinum* and other clostridia that produce botulinum neurotoxin, p. 3–20. *In* A. W. Hauschild and K. L. Dodds (ed.), *Clostridium botulinum*—ecology and control in foods. Marcel Dekker, New York.
19. Hauser, D., M. W. Eklund, P. Boquet, and M. R. Popoff. 1994. Organization of the botulinum neurotoxin C1 gene and its associated non-toxic protein genes in *Clostridium botulinum* C468. Mol. Gen. Genet. **243**:631–640.
20. Hutson, R. A., M. D. Collins, A. K. East, and D. E. Thompson. 1994. Nucleotide sequence of the gene coding for non-proteolytic *Clostridium botulinum* type B neurotoxin: comparison with other clostridial neurotoxins. Curr. Microbiol. **28**:101–110.
21. Hutson, R. A., M. D. Collins, Y. Zhou, J. Jiang, H. Sugiyama, E. A. Johnson, and C. L. Hatheway. 1996. Genetic characterization of *Clostridium botulinum* type A containing silent type B neurotoxin gene sequences. J. Biol. Chem. **271**:10786–10792.
22. Hutson, R. A., D. E. Thompson, P. A. Lawson, R. P. Schocken-Itturino, E. C. Bottger, and M. D. Collins. 1993. Genetic interrelationships of proteolytic *Clostridium botulinum* type A, B, and F and other members of the *Clostridium botulinum* complex as revealed by 16S rRNA gene sequences. Antonie van Leeuwenhoek **64**:278–283.
23. Lawson, P. A., S. E. Gharbia, H. N. Shah, and D. J. Clark. 1989. Recognition of *Fusobacterium nucleatum* subgroups Fn-1, Fn-2 and Fn-3 by ribosomal RNA gene restriction patterns. FEMS Microbiol. Lett. **65**:41–46.
24. Lonetto, M., M. Gribskov, and C. A. Gross. 1992. The σ^{70} family: sequence conservation and evolutionary relationships. J. Bacteriol. **174**:3843–3849.
25. Nukina, M., Y. Mochida, S. Sakaguchi, and G. Sakaguchi. 1991. Difficulties of molecular dissociation of *Clostridium botulinum* type G progenitor toxin. FEMS Microbiol. Lett. **79**:165–170.
26. Ohishi, I., and G. Sakaguchi. 1975. Molecular construction of *Clostridium*

botulinum type F progenitor toxin. Appl. Microbiol. **29:**444–447.

27. Saitou, N., and M. Nei. 1987. The neighbor-joining method: a new method for constructing phylogenic trees. Mol. Biol. Evol. **4:**406–425.

28. Somers, E., and B. R. DasGupta. 1991. *Clostridium botulinum* types A, B, C1, and E produce proteins with or without hemagglutinating activity: do they share common amino acids sequences and genes. J. Prot. Chem. **10:**415–425.

29. Sugii, S., I. Ohisi, and G. Sakaguchi. 1977. Correlation between oral toxicity and in vitro stability of *Clostridium botulinum* type A and B toxins of different molecular sizes. Infect. Immun. **16:**910–914.

30. Tabita, K., S. Sakaguchi, and G. Sakaguchi. 1991. Distinction between Clostridium botulinum type A strains associated with food-borne botulism and those with infant botulism in Japan in intraintestinal toxin production in infant mice and some other properties. FEMS Microbiol. Lett. **79:**251–256.

31. Thompson, D. E., J. K. Brehm, J. D. Oultram, T. J. Swinfield, C. C. Shone, T. Atkinson, J. Melling, and N. P. Minton. 1990. The complete amino acid sequence of the *Clostridium botulinum* type A neurotoxin deduced by nu-

cleotide sequence analysis of the encoding gene. Eur. J. Biochem. **189:**73–81.

32. Tsuzuki, K., K. Kimura, N. Fuji, N. Yokosawa, T. Indoh, T. Murakami, and K. Oguma. 1990. Cloning and complete nucleotide sequence of the gene for the main component of hemagglutinin produced by *Clostridium botulinum* type C. Infect. Immun. **58:**3173–3177.

33. Tsuzuki, K., K. Kimura, N. Fuji, N. Yokosawa, and K. Oguma. 1992. The complete nucleotide sequence of the gene coding for the nontoxic-nonhemagglutinin component of *Clostridium botulinum* type C progenitor toxin. Biochem. Biophys. Res. Commun. **183:**1273–1279.

34. Whelan, S. M., M. J. Elmore, N. J. Bodsworth, J. K. Brehm, T. Atkinson, and N. P. Minton. 1992. Molecular cloning of the *Clostridium botulinum* structural gene encoding the type B neurotoxin and determination of its entire nucleotide sequence. Appl. Environ. Microbiol. **58:**2345–2354.

35. Willems, A., A. K. East, P. A. Lawson, and M. D. Collins. 1993. Sequence of the gene coding for the neurotoxin of *Clostridium botulinum* type A associated with infant botulism: comparison with other clostridial neurotoxins. Res. Microbiol. **144:**547–556.

From:	Turner, Sharon
Sent:	Thursday, May 31, 2001 12:43 PM
To:	STIC-ILL
Subject:	08981087

Please provide

J. Clin. Microbiol., 31:2255-62, 1993

FEMS Microbiol., Lett. 108:175-182, 1993

Biochem. Biophys. Res Comm., 183:107-113, 1992

J. Gen. Microbiol., 137:519-525, 1991

Sharon L. Turner, Ph.D.
USPTO CM1-10B09
Biotechnology GAU 1647
(703) 308-0056

1

Cloning of a DNA fragment encoding the 5'-terminus of the botulinum type E toxin gene from *Clostridium butyricum* strain BL6340

Nobuhiro Fujii,[1] Kouichi Kimura,[1] Teruo Yashiki,[2] Tomokazu Indoh,[1] Tadayuki Murakami,[1] Kayo Tsuzuki,[1] Noriko Yokosawa[1] and Keiji Oguma[1]*

[1]*Department of Microbiology, Sapporo Medical College, Sapporo 060, Japan*
[2]*Laboratory of Technology, College of Medical Technology, Hokkaido University, Sapporo 060, Japan*

Chromosomal DNA was extracted from toxigenic *Clostridium butyricum* strain BL6340 isolated from a case of infant botulism. After digestion by *Eco*RI, a DNA fragment of about 1 kbp was cloned into *Escherichia coli* using λgt11, and was subcloned into pUC118. The *E. coli* cells transformed with this cloned fragment produced a 33 kDa protein which reacted with monoclonal antibodies recognizing the light chain (Lc) component of botulinum type E toxin. The nucleotide sequence of the cloned fragment was determined. The sequence was similar to that from botulinum type E toxin gene fragments previously determined by our laboratory (strains Mashike, Otaru and Iwanai). Several highly homologous sequences among the botulinum type A, C, E, butyricum and tetanus toxin genes were found in both translated and untranslated regions. These results suggest that the toxin gene of *C. butyricum* may have evolved by transfer from *C. botulinum*.

Introduction

Clostridium botulinum is the causative agent of food-borne botulism or infant botulism. Neurotoxins produced by the bacteria are classified into seven groups (A to G) based on their antigenicity. Recently, however, *Clostridium* species other than *C. botulinum* have been found to produce botulinum neurotoxin. *C. butyricum* and *C. barati* produce type E and type F neurotoxins, respectively, which are implicated in the onset of human infant botulism (Hall *et al.*, 1985; McCroskey *et al.*, 1986). The DNA extracted from a type E toxin-producing *C. butyricum* strain showed high homology to that obtained from a typical nontoxigenic *C. butyricum* strain (Suen *et al.*, 1988). It is therefore important to establish whether toxin genes exist in toxigenic and nontoxigenic *C. butyricum* strains, and to clarify the mechanism of transfer of toxin genes from *C. botulinum* to other clostridia. The N-terminal amino acid sequences of botulinum type A, B, C, D and E toxins, and of the toxin produced by toxigenic *C. butyricum*, have been determined (Gimenez *et al.*, 1988; Moriishi *et al.*, 1989;

Sathyamoorthy *et al.*, 1985; Tsuzuki *et al.*, 1988). The 5'-terminus and complete nucleotide sequence of type A toxin gene has been reported (Betlay *et al.*, 1989; Binz *et al.*, 1990; Thompson *et al.*, 1990). Recently, the complete nucleotide sequence of the structural gene of botulinum type C toxin and the 5'-terminal nucleotide sequence of the type E toxin gene (strains Mashike, Otaru, Iwanai) have been determined in our laboratory (Fujii *et al.*, 1990; Kimura *et al.*, 1990). The nucleotide sequences obtained from three type E toxin genes were identical. In this paper, we report the cloning, sequencing and expression of a toxin gene fragment from a toxigenic *C. butyricum* strain, and compare the 5'-terminal nucleotide sequence (and the deduced amino acid sequence) to those of botulinum toxins type A, C and E, and tetanus toxin.

Methods

Extraction of DNA from cells. One toxigenic strain of *C. butyricum* (BL6340, isolated from a case of infant botulism and kindly provided by Dr Hatheway, Centers for Disease Control, Atlanta, USA) and nine nontoxigenic strains (KZ145 and KZ589, provided by Dr Hatheway; IFO14810, IFO20064, Y-M-1a, Y-M-2a, Y-M-3a, Y-B-1a and IAM19001, provided by Dr Arai, Biotherapy Research Institute, Tokyo, Japan) were used. These strains were cultured in 250 ml LYG medium (1% (w/v) lactalbumin hydrolysate, 2% (w/v) yeast extract,

Abbreviation: mAb, monoclonal antibody.

0·5% (w/v) glucose, 0·15% (w/v) cysteine hydrochloride, pH 7·2) at 30 °C for 12 h. The cells were collected by centrifugation at 6000 r.p.m. (11000 g) for 10 min at room temperature, and resuspended in 20 ml LYG medium containing 8% (w/v) polyethylene glycol 4000, 10 U penicillin G ml⁻¹ and 10 μg lysozyme ml⁻¹, and then incubated at 30 °C for 60–90 min. After centrifugation, the cells were lysed using 2–3% (w/v) of sodium dodecyl sulphate (SDS). Extraction and purification of chromosomal DNA were performed by the procedures described by Marmur (Marmur, 1961).

Monoclonal antibodies. Three monoclonal antibodies (mAbs) against light chain (Lc) component of type E toxin, EL161-38, EL211-3, EL219-15, were those prepared previously (Tsuzuki *et al.*, 1988). EL211-3 and EL219-15 react only with type E toxin, but EL161-38 also reacts with Lc components of botulinum type B, C, D and tetanus toxins.

Cloning and sub-cloning of the toxin gene. Cellular DNA from toxigenic strain BL6340 was digested by *Eco*RI (Takara Shuzo Co., Kyoto, Japan), ligated to λgt11 *Eco*RI arms (Promega Corporation), and then packed *in vitro* by using Packagene (Promega Corporation). The gene library was plated on *E. coli* Y1090 (Promega Corporation), and positive clones were immunoscreened using mAb EL161-38 and the ProtoBlot Immunoscreening System (Promega Corporation). The isolated positive clones were purified, and the recombinant phage DNA was extracted. After digestion with *Eco*RI, the inserts separated from phage DNA were then sub-cloned into the *Eco*RI site on the polylinker of pUC118 (Takara Shuzo Co.).

DNA sequencing and Western blot analysis. DNA sequence analysis was performed by the dideoxy termination method using [α-³⁵S]dATP (NEN Products) and a T7 DNA sequencing kit (Pharmacia). Gene products from *E. coli* strain Y1090 infected with recombinant phage, and from *E. coli* strain MV1184 transformed by recombinant plasmids, were analysed by Western blots with mAbs and with alkaline-phosphatase-labelled anti-mouse immunoglobulin G conjugate (Promega Corporation), nitroblue tetrazolium and 5-bromo-4-chloro-3-indolyl phosphate. The transformed cells were incubated in 10 ml 2×TY medium [1·6% (w/v) tryptone, 1% (w/v) yeast extract, 0·5% (w/v) NaCl, pH 7·6] at 37 °C for 12 h. After centrifugation, the cells were suspended in 1 ml TEP buffer (100 mM-Tris/HCl, pH 7·4, 10 mM-EDTA, 1 mM-PMSF), sonicated for about 2 min at 0 °C, and centrifuged at 12000 r.p.m. for 10 min at 4 °C. The resulting supernatants were electrophoresed on 12% (w/v) SDS-polyacrylamide gel, and Western blot analysis was carried out as described previously (Tsuzuki *et al.*, 1988).

Dot blot hybridization. DNA samples (2, 20 and 40 μg) from toxigenic and nontoxigenic strains of *C. butyricum* were spotted on a nylon membrane (Pall Ultrafine Filtration Corp., Glen Cove, NY, USA). After denaturation and neutralization, the membrane was prehybridized in 50% (v/v) formamide hybridization buffer at 42 °C for 12 h, and then hybridized with ³⁵S-labelled toxin probe using a DNA labelling kit (Nippon Gene, Toyama, Japan) at 42 °C for 24 h according to the procedures described previously (Fujii *et al.*, 1988). The membrane was washed at 68 °C and autoradiographed.

Results and Discussion

Cloning and gene products

Chromosomal DNA isolated from the toxigenic *C. butyricum* strain BL 6340 was digested with *Eco*RI and ligated to λgt11 *Eco*RI arms. After *in vitro* packaging,
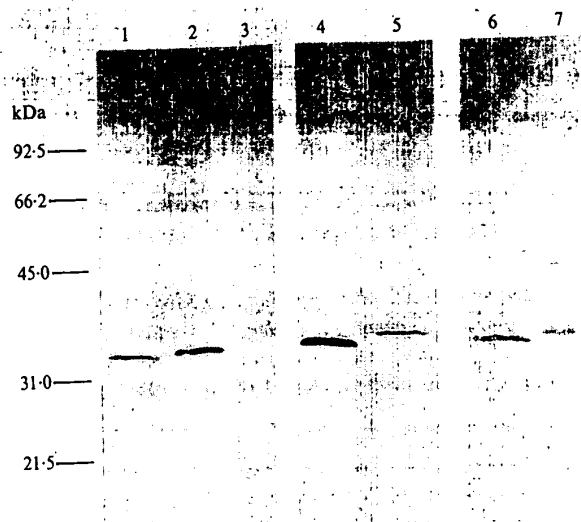
Fig. 1. Western blot analysis of gene products by recombinant phage and plasmid. Cellular extracts from *E. coli* Y1090 infected with (1) λgt11 containing a 1 kbp *Eco*RI fragment of the butyricum toxin gene (λBU-L1); (2) λgt11 containing a type E toxin gene *Eco*RI fragment (λEM-L1); (3) control λgt11; and from *E. coli* MV1184 transformed by pUC118 containing butyricum toxin gene *Eco*RI fragment (pBU-L1) (4 and 6) or botulinum type E toxin gene *Eco*RI fragment (pEM-L1) (5 and 7). The extracts were characterized by Western blots using mAb EL161-38 against botulinum type E toxin light chain. The cells of lanes 1–5, and lanes 6–7 were incubated in the presence and absence of IPTG, respectively.

and infection into *E. coli*, the recombinant plaques were screened for the production of toxin fragments using mAb EL161-38. Several positive clones were detected and one, λBU-L1, was chosen for further study. λBU-L1 DNA was prepared, and found to contain a single 1 kbp *Eco*RI insert. *E. coli* Y1090 cells infected with phage λBU-L1 produced a 33 kDa protein which reacted with mAb EL161-38 both in the presence and in the absence of IPTG (Fig. 1). In addition, the infected cells reacted with mAb EL211-3, but not with mAb EL219-45 (data not shown). The 1·0 kbp *Eco*RI fragment was isolated from λBU-L1, and then recloned into the *Eco*RI site of pUC118. *E. coli* strain MV1184, when transformed by the recombinant plasmid (pBU-L1) produced a protein of similar size (34 kDa) as that obtained from the phage clone (Fig. 1). These gene products migrated slightly faster than those produced by *E. coli* cells that had been transformed with phage or plasmid expressing an analogous fragment (about 1 kbp) which encodes the botulinum type E toxin gene (Fujii *et al.*, 1990).

Susceptibility of insert to endonuclease

The 1·0 kbp *Eco*RI fragment from pBU-L1 was digested with several restriction enzymes, and the restriction map
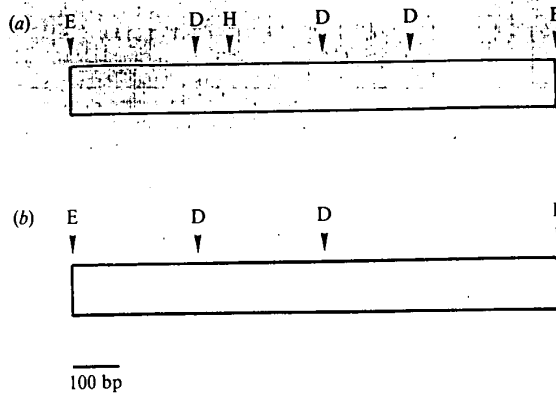
Fig. 2. Restriction endonuclease recognition sites in the fragments from *C. butyricum* (*a*) and *C. botulinum* type E (*b*). The map was determined by 4% NuSieve agarose gel electrophoresis of the plasmid DNAs cleaved by *Eco*RI (E), *Dra*I (D) and *Hinc*II (H).

was found to be similar, but not identical, to that of the analogous 1·0 kbp fragment from the *C. botulinum* type E toxin gene (Fig. 2).

Nucleotide and derived amino acid sequences of butyricum toxin gene EcoRI fragment

The 5'-terminal nucleotide sequence (756 bp) of the butyricum toxin gene and 227 bp of the 5'-untranslated region was determined. The number of nucleotides of the coding and noncoding regions was the same as that of the cloned type E toxin gene fragment (Fujii *et al.*, 1990). However, the nucleotide sequences of the butyricum and botulinum type E toxin genes were not identical; eight bases in the untranslated region, and 27 bases in the open reading frame were different (Fig. 3). A Shine–Dalgarno consensus sequence (AGGAGA) was found 14 nucleotides upstream from the translation initiation start codon (ATG). A base sequence corresponding to the −10 consensus region of *E. coli* (TATAAT) was identified at position 87, but no sequence corresponding to the −35 consensus sequence of *E. coli* (TTGACA) was detected. Two stem-loop structures, at bases 17 to 48 and 121 to 155, were also detected in the region upstream of the −10 region and the Shine–Dalgarno sequence, respectively. The deduced N-terminal amino acid sequence of butyricum toxin was identical to that previously reported by protein analysis of the toxin (Gimenez *et al.*, 1988). It was also similar to the protein sequence of the botulinum type E toxin (strains Mashike, Otaru and Iwanai) except for 17 amino acid residues, which were altered by the nucleotide changes described above (Fig. 4).

Comparison of nucleotide and amino acid sequences of the putative toxin gene with those of botulinum type A, C and tetanus toxins

As described above, the nucleotide sequences of butyricum and type E toxin gene fragments were similar in both the noncoding and coding regions. Since the nucleotide sequences of types A, C and E, and tetanus toxin gene have been determined, the sequences of these five toxin genes were compared. There exist several regions of homologous sequence in both untranslated and translated regions (Fig. 3). In the untranslated region, it was very easy to find sequences corresponding to a Shine–Dalgarno sequence which is presumably involved in the initiation of translation of the toxin genes. In contrast, the sequence corresponding to the −10 and −35 regions was difficult to identify, due to the high A-T content of the DNA. However, one potential −10 region was identified 121 bp upstream from a Shine–Dalgarno region.

In addition to these probable −10 and Shine–Dalgarno regions, we identified highly homologous sequences of ATGATAAAT and TATAAAAAATCA at about 2 and 21 bp downstream of the probable −10 region of the *C. butyricum*, *C. botulinum* type E and tetanus toxin genes, and the sequence ATATTAAA was found between the probable promoter and Shine–Dalgarno regions in all of the *C. butyricum*, type A, type C, type E and tetanus toxin genes. We propose that these homologous sequences may play a role in regulation of expression of the toxin genes.

It has been reported that tetanus toxin and the different types of botulinum toxins possess several consensus amino acid sequences of between 7 and 20 residues in length (Tsuzuki *et al.*, 1988). In addition to these sequences, we found three other highly homologous regions in botulinum type A, C, E, butyricum and tetanus toxins (Fig. 4). The nucleotide sequences corresponding to these three homologous regions were also quite similar to each other (Fig. 3). One region (amino acid residues 65 to 85 in butyricum toxin) was hydrophilic, and another, the region 201–223, was hydrophobic. This hydrophobic region might conceivably play a role in spanning the mammalian cell membrane, as suggested for tetanus toxin by Eisel *et al.* (1986).

Detection of toxin gene in toxigenic and nontoxigenic strains of C. butyricum

DNA samples (2, 20, and 40 µg) from one toxigenic and nine nontoxigenic strains of *C. butyricum* were spotted on a nylon membrane, and the membrane was hybridized with the 1 kbp *Eco*RI probe from the *C. butyricum* toxin

70

Butyricum (BL6340)   GAATTCAATT AGTAGATAAC AAAAATAATG CAAAGACTTT TATTATTAAT AATGATATAT TTATTTCTAA
                                        ←―――――――――――― -5.6 ――――――――→

Type E (Mashike)   ---------C- ----------T ---------- --C---T--- ---------- ---------- ----C-----

Type A (62A)   ACTAATAGAT AACAAAAATA ACGCAAAGAA GATGATAATT AGTAATGATA TATTTATTTC CAATTGTTTA

Type C (c·st)   TCCTGAAAAT AATAGAATAC AATTAGTAAG TTCCAAAGAT AATGCAAAAA AGATTACAGT TAATACTGAT

Tetanus   TTATAATTTA ATTATGAATA ATATTCTTAA GATAAAAAGT AAATTTTTAA AAATTTAAAT TTTCAGTTTA
                                                                                          •

140

Butyricum   CTGTTTAACT TTAACTTATA ATAATGTAAA TGTATATTTA TCTATAAAAA ATCAAGATTA CAATTGGGTT
                                        ――――――   ――――――――――   ←――――――――――――――― -6.3

Type E   ---------- ---------- -C-------- ---------G ---------- ---------- ----------

Type A   ACCCTATCTT ATAACGGTAA ATATATATGT TTATCTATGA AAGATGAAAA CCATAATTGG ATGATATGTA

Type C   TTATTTAGAC CTGATTGTAT AACATTTTCA TATAATGATA AATATTTTTC TCTATCACTT ACAGATGGAG

Tetanus   CAAAAAATAA CCTGATTATG TTATATGTAA TTGTAAAAAA CATATAAAAA ATCAGAAAAA TTTAGGAGGT
                                        ――――――――   ――――――――

210

Butyricum   ATATGTGATC TTAATCATGA TATACCCAAA AAGTCATATC TATGGATATT AAAAAATATA TAAATTTAAA
                      ――――――――――――――――――――――――――――――→

Type E   ---------- ---------- ------A--- ---------- ---------- ---------- ----------

Type A   ATAATGATAT GTCAAAGTAT TTGTATTTAT GGTCATTTAA ATAATTAATA ATTTAATTAA TTTTAAATAT

Type C   ATTATAATTG GATGATATGT AATGACAATA ACAAGGTGCT AAACGTGCAC ATTTGTGGAT ATTAGAAAGT

Tetanus   ATATTATTAA TGGATTAAAT AATAATTTTT TAATTTACTT TTGATTAATA AATATTAAAT GTTTATTTTA
                                                                            ――――――――――――

280

                                              ┌→ ORF

Butyricum   ATTAGGAGAT GCTGTATATG CCAACAATTA ATAGTTTTAA TTATAATGAT CCTGTTAATA ATAGAACAAT
                         │┌─────┐│   │
Type E   ---│------│- ------│--- --- ----A----- ---------- ---------- ---------C ----------

Type A   TATAAGAGGT GTTAAATATG CCATTTGTTA ATAAACAATT TAATTATAAA GATCCTGTAA ATGGTGTTGA

Type C   TAGGAGATGT TAGTATTATG CCAATAACAA TTAACAACTT TAATTATTCA GATCCTGTTG ATAATAAAAA

Tetanus   ATTAGGAGAT GATACGTATG CCAATAACCA TAAATAATTT TAGATATAGT GATCCTGTTA ATAATGATAC
                          SD

350

Butyricum   TTTATATATT AAACCAGGCG GTTGTCAACA ATTTTATAAA TCATTTAATA TTATCAAAAA TATTTCGATA

Type E   ---------- ---------- ---------G- ---------- ---------- ---------- ----------

Type A   TATTGCTTAT ATAAAAATTC CAAATGCAGG ACAAATGCAA CCACTAAAAG CTTTTAAAAT TCATAATAAA

Type C   TATTTTATAT TTAGATACTC ATTTAAATAC ACTAGCTAAT GAGCCTGAAA AAGCCTTTCG CATTACAGGA

Tetanus   AATTATTATG ATGGAGCCAC CATACTGTAA CGGTCTAGAT ATCTATTATA AGGCTTTCAA AATAACAGAT
                                              ――――――――――――

420

Butyricum   ATTCCAGAGA GAAATGTAAT TGGTACAATT CCCCAAGATT TTCTTCCGCC TACTTCATTG AAAAATCGAG

Type E   ---------- ---------- --------CC ---------- ---A------ ---------A ----------

Type A   ATATGGGTTA TTCCAGAAAG AGATACATTT ACAAATCCTG AAGAAGGAGA TTTAAATCCA CCACCAGAAG

Type C   AATATATCGG TAATACCTGA TAGATTTTCA AGAAATTCTA ATCCAAATTT AAATAAACCT CCTCCAGTTA

Tetanus   CGTATTTGGA TAGTGCCGGA AACGTATGAA TTTGGGACAA AACCTGAAGA TTTTAACCCA CCATCTTCAT

490

                      ┌─ C   ┌─ A, T

Butyricum   ATAGTAGTTA TTATGACCCT AATTATTTAC AAAGTGATCA AGAAAAGGAT AAATTTTTAA AAATAGTCAC

Type E   ---------- ---------- ---------- --------C- ---------- -G-------- ----------

Type A   CAAAACAAGT TCCAGTTTCA TATTATGATT CAACATATTT AAGTACAGAT AATGAAAAAG ATAATTATTT

Type C   CAAGCCCTAA AAGTGGTTAT TATGATCCTA ATTATTTGAG TACTGATTCT GACAAAGATA CATTTTTAAA

Tetanus   TAATAGAAGG TGCATCTGAG TATTACCATC CAAATTATTT AAGGACTGAT TCTGATAAAG ATACATTTTT

560

Butyricum   AAAAATATTT AATAGAATAA ATGATAATCT TTCAGGAAGG ATTTTATTAG AAGAACTGTC AAAAGCTAAT

Type E   ---------- ---------- --A------- -------G-- ---------- ---------- ----------

```
Type A      AAAGGGAGTT ACAAAATTAT TTGAGAGAAT TTATTCAACT GATCTTGGAA GAATGTTGTT AACATCAATA

Type C      AGAAATTATA AAGTTATTTA AAAGAATTAA TTCTAGAGAA ATAGGAGAAG AATTAATATA TAGACTTTCG

Tetanus     ACAAACCATG GTAAAACTGT TTAACAGAAT TAAAAACAAT GTAGCAGGTG AAGCCTTATT AGATAAGATA

                                                                                      630
Butyricum   CCATATTTAG GAAATGATAA TACTCCAGAT GGTGACTTCA TTATTAATGA TGCATCAGCA GTTCCAATTC

Type E      ---------- -G-------- ---------- AA-C-A---C A----GG--- ---------- ---GAG---A


Type A      GTAACGGGAA TACCATTTTG GGGTGGAAGT ACAATAGATA CAGAATTAAA AGTTATTGAT ACTAATTGTA

Type C      ACAGATATAC CCTTTCCTGG GAATAACAAT ACTCCAATTA ATACTTTTCA TTTTGATGTA GATTTTAACA

Tetanus     ATAAATGCCA TACCTTACCT TGGAAATTCA TATTCCTTAC TAGACAAGTT TGATACAAAC TCTAATTCAG

                                                                                      700
Butyricum   AATTCTCAAA TGGTAGCCAA AGCATACTAT TACCTAATGT TATTATAATG GCAGCAGAGC CTGATTTATT

Type E      ---------- ---------- GA-------- ---------- ---------- ---------- ----------


Type A      TTAATGTGAT ACAACCAGAT GGTAGTTATA GATCAGAAGA ACTTAATCTA GTAATAATAG GACCCTCAGC

Type C      GTGTTGATGT TAAAACTAGA CAAGGTAACA ACTGGGTTAA AACTGGTAGC ATAAATCCTA GTGTTATAAT

Tetanus     TATCTTTTAA TTTATTAGAA CAAGACCCCA GTGCAGCAAC TACAAAATCA GCAATGCTGA CAAATTTAAT

                                                                                      770
Butyricum   TGAAACTAAC AGTTCCAATA TTTCTCTAAG AAATAATTAT ATGCCAAGCA ATCACGGTTT TGGATCAATA

Type E      ---------- ---------- ---------- ---------- ---------- ---------- ----------


Type A      TGATATTATA CACTTTGAAT GTAAAAGCTT TGGACATGAA GTTTTGAATC TTACGGGAAA TGGTTATGGC

Type C      AACTGGACCT AGAGAAAACA TTATAGATCC AGAAACTTCT ACGTTTAAAT TAACTAACAA TACTTTTCCG

Tetanus     AATATTTGGA CCTGGGCCTG TTTTAAATAA AAATGAGGTT AGAGGTATTG TATTGAGGGT AGATAATAAA
                                                            ┌─C        ┌─T   840
Butyricum   GCTATAGTAA CATTCTCACC TCAATATTCT TTTACATTTA AAGATAATAG TATCAATCAA TTTATTCAAG

Type E      ---------- ---------- ---------- ---------- -T-------- ---------- ----------


Type A      TCTACTCAAT ACATTAGATT TAGCCCAGAT TTTACATTTG GTTTTGAGGA GTCACTTCAA GTTGATACAA

Type C      GCACAAGAAG GATTTGGTGC TTTATCAATA ATTTCAATAT CACCTAGATT TATGCTAACA TATAGTAATG

Tetanus     AATTACTTCC GATGTAGAGA TGGTTTTGGC TCAATAATGC AAATGGCATT TTGCCCAGAA TATGTACCTA
                                                                                      910
                          ┌─A
Butyricum   ATCCTGCTCT TACATTAATG GATGAATTAA TACATTCATT ACATGGACTA TATCGGGCTA AAGGGCATTAC

Type E      ---------- ---------- ---------- ---------- ---------- ---------- ----------


Type A      ATCCTCTTTT ACGTGCAGGC AAATTTGCTA CACATCCACC AGTAACATTA GCACATGAAC TTATACATGG

Type C      CAACTAATGA TGTAGGAGAG GGTAGCATTT CTAAGTCTGA ATTTTGCATG GATCCAATAC TAATTTTAAT

Tetanus     CCTTTGATAA TGTAATAGAA AATATTACGT CACTCACTAT TGGCAAAAGC AAATATTTTC AAGATCCAGC

                                                                                      983
Butyricum   TACAATGTAT ACTATAACAC AAAAACAAAA TCCCCTAATA ACAAATATAA GAGGTACAAA TATTGAAGAA TTC

Type E      ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---


Type A      TGGACATAGA TTATATGGAA TAGCAATTAA TCCAAATAGG GTTTTTAAAG TAAATACTAA TGCCTATTAT GAA

Type C      GCATGAACTT AATCATGCAA TGCATAATTT ATATGGAATA GCTATACCAA ATGATGAAAC AATTTCATCT GTA

Tetanus     ATTACTATTA ATGCACGAAC TTATACATGT ACTACATGGT TTATAGGAA TGCAGGTATC AAGCCATGAA ATT
```

Fig. 3. Nucleotide sequences of the 5′-translated and 5′-untranslated regions of butyricum, botulinum and tetanus toxin genes. The DNA sequence of the butyricum toxin gene *Eco*RI fragment is compared to the sequence of botulinum type A (Binz *et al.*, 1990), type C (Kimura *et al.*, 1990), type E (Fujii *et al.*, 1990) and tetanus (Eisel *et al.*, 1986) toxin genes. The putative Shine–Dalgarno sequence is indicated by SD. The sequences homologous to the −10 region of the *E. coli* promoter (TATAAT) are marked with double underlines. A sequence with similar dyad symmetry is underlined by divergent arrows. Numbers represent free energy values in kcal mol^{-1} (1 kcal = 4·2 kJ). Dashes in the botulinum type E sequence represent bases identical to the *C. butyricum* sequence. The highly homologous sequences are underlined, and the start site of homology is indicated on the butyricum sequence by A (type A), C (type C) and T (tetanus). ORF is the translation initiation site of ATG.

```
                    ↑↑oooo  ↑o↑  ↑↑↑oo      ↑o                                    ↑↑  40
Butyricum (BL6340)  M P T I N S F N Y N  D P V N N R T I L Y  I K P G G C Q Q F Y  K S F N I M K N I W
Type E (Mashike)    - - K - - - - - - -  - - - D - - - -      - - - - - - - E - -  - - - - - - - - - -

Type A (62A)        M P F V N K Q F N Y  K D P V N G V D I A  Y I K I P N A G Q M  Q P V K A F K I H N
Type C (c-st)       M P I T I N N F N Y  S D P V D N K N I L  Y L D T H L N T L A  N E P E K A F R I T
Tetanus             M P I T I N N F R Y  S D P V N N D T I I  M M E P P Y C K G L  D I Y Y K A F K I T


                    o↑o↑                                     ↑↑↑  oo↑↑   ↑    ↑  80
Butyricum           I I P E R N V I G T  I P Q D F L P P T S  L K N G D S S Y Y D  P N Y L Q S D Q E K
Type E              - - - - - - - - - -  T - - - - H - - - -  - - - - - - - - - -  - - - - - - - E - -

Type A              K I W V I P E R D T  F T N P E E G D L N  P P P E A K Q V P V  S Y Y D S T Y L S T
Type C              G N I W V I P D R F  S R N S N P N L N K  P P R V T S P K S G  Y Y D P N Y L S T D
Tetanus             D R I W I V P E R Y  E F G T K P E D F N  P P S S L I E G A S  E Y Y D P N Y L R T


                    ↑  o↑o                                                        120
Butyricum           D K F L K I V T K I  F N R I N D N L S G  R I L L E E L S K A  N P Y L G N D N T P
Type E              - R - - - - - - - -  - - - - N - - - - G  - - - - - - - - - -  - - - - - - - - - -

Type A              D N E K D N Y L K G  V T K L F E R I Y S  T D L G R M L L T S  I V R G I P F W G G
Type C              S D K D T F L K E I  I K L F K R I N S R  E I G E E L I Y R L  S T D I P F P G N N
Tetanus             D S D K D R F L Q T  M V K L F N R I K N  N V A G E A L L D K  I I N A I P Y L G N


                                                                                  160
Butyricum           D G D F I I N D A S  A V P I Q F S N G S  Q S I L L P N V I I  M G A E P D L F E T
Type E              - N Q - H - G - - -  - - E - K - - - - -  - D - - - - - - - -  - - - - - - - - G -

Type A              S T I D T E L K V I  D T N C I N V I Q P  D G S Y R S E E L N  L V I I G P S A D I
Type C              N T P I N T F D F D  V D F N S V D V K T  R Q G N N W V K T G  S I N P S V I I T G
Tetanus             S Y S L L D K F D T  N S N S V S F N L L  E Q D P S G A T T K  S A M L T N L I I F


                                                                                  200
Butyricum           N S S N I S L R N N  Y M P S N H G F G S  I A I V T F S P E Y  S F R F K D N S M N
Type E              - - - - - - - - - -  - - - - - - - - - -  - - - - - - - - - -  - - - - N - - - - -

Type A              I Q F E C K S F G H  E V L N L T R N G Y  G S T Q Y I R F S P  D F T F G F E E S L
Type C              P R E N I I D P E T  S T F K L T N N T F  A A Q E G F G A L S  I I S I S P' R F M L
Tetanus             G P G P V L N K N E  V R G I V L R V D N  K N Y F P C R D G F  G S I M Q H A F C P


                    ↑↑oo  ↑ o↑↑↑o↑  ↑  ↑↑↑                                       240
Butyricum           E F I Q D P A L T L  M H E L I H S L H G  L Y G A K G I T T M  Y T I T Q K Q N P L
Type E              - - - - - - - - - -  - - - - - - - - - -  - - - - - - - - - -  - - - - - - - - - -

Type A              E V D T N P L L G A  G K F A T D P A V T  L A H E L I H A G H  R L Y G I A I N P N
Type C              T Y S N A T N D V G  E G R F S K S E F C  M D P I L I L M H E  L N H A M H N L Y G
Tetanus             E Y V P T F D N V I  E N I T S L T I G K  S K Y F Q D P A L L  L M H E L I H V L H


Butyricum           I T N I R G T N I E  E F
Type E              - - - - - - - - - -  - -

Type A              R V F K V N T N A Y  Y E
Type C              I A I P N D Q T I S  S V
Tetanus             G L Y G M Q V S S H  E I
```

Fig. 4. Amino acid sequences derived from the base sequence of butyricum toxin gene *Eco*RI fragment. Dashes in the botulinum type E sequence represent amino acid residues identical to the *C. butyricum* sequence. The highly conserved homologous sequences are underlined. Arrows indicate highly conserved amino acid residues among the five toxins, and open circles indicate common amino acid residues in four out of five toxins.

gene. A positive reaction was observed only with DNA from the toxigenic strain, BL6340 (data not shown). The nucleotide sequences of the toxin genes from *C. butyricum* and *C. botulinum* strains were quite similar in both coding and noncoding regions. Therefore, it is quite possible that a nontoxigenic *C. butyricum* strain has become toxigenic by acquiring the toxin gene from a *C. botulinum* type E strain; the toxin gene could perhaps have been transmitted from *C. botulinum* type E to *C. butyricum* cells through some transmissible vector, e.g. a bacteriophage or a plasmid.

References

BETLAY, M. J., SOMERS, E. & DASGUPTA, B. R. (1989). Characterization of botulinum type A neurotoxin gene: Delineation of the N-terminal encoding region. *Biochemical and Biophysical Research Communications* **162**, 1388–1395.

BINZ, T., KURAZONO, H., WILLE, M., FREVERT, J., WERNARS, K. & NIEMANN, H. (1990). The complete sequence of botulinum neurotoxin type A and comparison with other clostridial neurotoxins. *Journal of Biological Chemistry* **256**, 9153–9158.

EISEL, U., JARAUSCH, W., GORETZKI, K., HENSCHEN, A., ENGELS, J., WELLER, U., HUDEL, M., HAKERMANN, E. & NIEMANN, H. (1986). Tetanus toxin: primary structure, expression in *E. coli*, and homology with botulinum toxins. *EMBO Journal* **5**, 2495–2502.

FUJII, N., OGUMA, K., YOKOSAWA, N., KIMURA, K. & TSUZUKI, K. (1988). Characterization of bacteriophage nucleic acids obtained from *Clostridium botulinum* types C and D. *Applied and Environmental Microbiology* **54**, 69–73.

FUJII, N., KIMURA, K., INDOH, T., MURAKAMI, T., YASHIKI, T., TSUZUKI, K., YOKOSAWA, N. & OGUMA, K. (1990). The nucleotide and deduced amino acid sequences of *Eco*RI fragment containing the 5′-terminal region of *Clostridium botulinum* type E toxin gene

cloned from Mashike, Iwanai and Otaru strains. *Microbiology and Immunology* **34**, 1041–1047.

GIMENEZ, J., FOLEY, J. & DASGUPTA, B. R. (1988). Neurotoxin type E from *Clostridium botulinum* and *C. butyricum*; Partial sequence and comparison. *FASEB Journal* **2**, A1750(8447).

HALL, J. D., MCCROSKEY, L., PINCOMB, B. J. & HATHEWAY, C. L. (1985). Isolation of an organism resembling *Clostridium barati* which produces type F botulinal toxin from an infant with botulism. *Journal of Clinical Microbiology* **21**, 654–655.

KIMURA, K., FUJII, N., TSUZUKI, K., MURAKAMI, T., INDOH, T., YOKOSAWA, N. & OGUMA, K. (1990). The complete nucleotide sequence of the gene coding for botulinum type C_1 toxin in the c-st phage genome. *Biochemical and Biophysical Research Communications* **171**, 1304–1311.

MCCROSKEY, L. M., HATHEWAY, C. L., FENICIA, L., PASOLINI, B. & AURELI, P. (1986). Characterization of an organism that produces type E botulinal toxin but which resembles *Clostridium butyricum* from the feces of an infant with type E botulism. *Journal of Clinical Microbiology* **23**, 201–202.

MARMUR, J. (1961). A procedure for the isolation of deoxyribonucleic acid from micro-organisms. *Journal of Molecular Biology* **3**, 208–218.

MORIISHI, K., SYUTO, B., KUBO, S. & OGUMA, K. (1989). Molecular diversity of neurotoxins from *Clostridium botulinum* type D strains. *Infection and Immunity* **57**, 2886–2891.

SATHYAMOORTHY, V. & DASGUPTA, B. R. (1985). Separation, purification, partial characterization and comparison of heavy and light chains of botulinum neurotoxin types A, B, and E. *Journal of Biological Chemistry* **260**, 10461–10466.

SUEN, J. C., HATHEWAY, C. L., STEIGERWALT, A. G. & BRENNER, D. J. (1988). Genetic confirmation of identities of neurotoxigenic *Clostridium barati* and *Clostridium butyricum* implicated as agents of infant botulism. *Journal of Clinical Microbiology* **26**, 2191–2192.

THOMPSON, E. D., BREHM, K. J., OULTRAM, D. J., SWINFIELD, T. J., SHONE, C. C., ATKINSON, T., MELLING, J. & MINTON, P. N. (1990). The complete amino acid sequence of the *Clostridium botulinum* type A neurotoxin, deduced by nucleotide sequence analysis of the encoding gene. *European Journal of Biochemistry* **189**, 73–81.

TSUZUKI, K., YOKOSAWA, N., SYUTO, B., OHISHI, I., FUJII, N., KIMURA, K. & OGUMA, K. (1988). Establishment of a monoclonal antibody recognizing an antigenic site common to *Clostridium botulinum* type B, C_1, D and E toxins and tetanus toxin. *Infection and Immunity* **56**, 898–902.

STIC-ILL

From: Turner, Sharon
Sent: Thursday, May 31, 2001 4:46 PM
To: STIC-ILL
Subject: 08981087

BIOCHIMIE, (1989 Nov-Dec) 71 (11-12) 1193-200.

BIOCHEMICAL JOURNAL, (1990 May 15) 268 (1) 123-8.

INFECTION AND IMMUNITY, (1992 Feb) 60 (2) 518-22.

Sharon L. Turner, Ph.D.
USPTO CM1-10B09
Biotechnology GAU 1647
(703) 308-0056

C. *Botulinum* neurotoxin types A and E: isolated light chain breaks down into two fragments. Comparison of their amino acid sequences with tetanus neurotoxin

Bibhuti R. DASGUPTA and John FOLEY Jr.

Food Research Institute, University of Wisconsin, 1925 Willow Drive, Madison, WI 53706, USA

(Received 15 June 1989, accepted 23 August 1989)

Summary — The flaccid paralysis in the neuromuscular disease botulism appears to depend on the coordinated roles of the ~50 kDa light and ~100 kDa heavy chain subunits of the ~150 kDa neurotoxic protein produced by *Clostridium botulinum* (*J. Biol. Chem.* (1987) 262, 2660 and *Eur. J. Biochem.* (1988) 177, 683). We observed that the light chain after separation from its conjugate heavy chain, in the presence of dithiothreitol and 2 M urea, begins to split into ~28 and ~18 kDa fragments. The other subunit–the ~100 kDa heavy chain following its isolation–and the parent ~150 kDa dichain neurotoxin do not break down under comparable conditions. This cleavage was examined in the neurotoxin serotypes A and E. The cleavage does not appear to be due to a protease. Partial amino acid sequences established that: i) the ~28-kDa and ~18-kDa fragments comprise the N- and C-terminal regions of the light chain, respectively; ii) the light chain of the neurotoxin serotypes A and E break down at precise peptide bonds; iii) the peptide bonds cleaved in serotypes A and E are five residues apart; and iv) the portions of the ~18 kDa fragments of serotype A and E neurotoxin sequenced so far are highly homologous to the corresponding region of tetanus neurotoxin produced by *Clostridium tetani*. The partial N-terminal sequence of the ~28 kDa fragment matches with the N-terminal sequence of the intact L chain. The 47 residues of the ~18-kDa fragment of type A sequenced from its N-terminal are: – *Y*.E.M.S.G.L.E.V.*S*.F.*E*.*E*.*L*.R.*T*.*F*.*G*.*G*.H.*D*.*A*.K.F.*I*.D.S.L.Q.E.*N*.E.F.R.L.Y.Y.Y.*N*.K.F.*K*. D.*I*.*A*.S.T.*L*. –. These align with those of tetanus neurotoxin beginning at its residue #259 (Tyr); the 18 underlined residues of the above 47 residues (*i.e.* 38%) are identical in positions between the two proteins. The 41 residues sequenced from the ~18 kDa fragment of type E *botulinum* neurotoxin are: – K.G.*I*.N.I.*E*.*E*.F.L.   *T*.*F*.*G*.N.N.*D*.L.*N*.I.*I*.T.V.*A*.Q.*Y*.*N*.*D*.I.*Y*.T.N.L.*L*.*N*.*D*.*Y*.R.*K*.*I*.*A*.X.K. *L*. –. These align with tetanus neurotoxin, beginning at its residue #264 (Tyr); the 20 underlined residues (*i.e.* 49%) are positionally identical to tetanus neurotoxin. Also the 13 overlined residues of type E (*i.e.* 32%) are positionally identical to type A *botulinum* neurotoxin.

***botulinum* type A and E neurotoxin / light chain / tetanus neurotoxin / sequence / homology**

Introduction

Botulinum neurotoxin (NT) blocks release of the excitatory neurotransmitter acetylcholine, and thus causes flaccid paralysis in the neuromuscular disease, botulism. The ~150 kDa NT is produced by *Clostridium botulinum* in 7 anti-genically distinct forms, called types A–G. The single chain polypeptide is nicked at 1/3 the distance from the N-terminal by protease endogenous to the bacteria or by exogenous proteases such as trypsin. The ~50 kDa light (or L) and ~100 kDa heavy (or H) chain subunits, containing the original N- and C-terminal ends,

respectively, remain linked by noncovalent and disulfide bond(s). The two subunits have been separated chromatographically and analyzed for amino acid compositions and partial amino acid sequences [1–4]. These two subunits of the NT appear to have distinct roles in the mechanism of action of the ~150 kDa NT. Experiments with neuromuscular junction preparations and isolated L and H subunit chains from types A and B NTs suggest that binding of the H chain with the specific sites on the nerve terminals permits the L chain or some combination of the L and H chains to enter the cell; then presumably the L chain causes blockade of release of the neurotransmitter acetylcholine and thus brings about paralysis [5, 6]. Recently types A and B NTs and the L chains isolated therefrom, but not the corresponding isolated H chains, were found to inhibit catecholamine release from digitonin-permeabilized bovine chromaffin cells [7].

We noticed [4] that the L chain of type A NT, following its separation from the H chain and storage for some time, yielded by polyacryl-amide gel electrophoresis in the presence of SDS (PAGE–SDS) two extra bands, ~28 and ~18 kDa, in addition to the 53 kDa L chain band. These or similar bands did not originate band. These or similar bands did not originate from the separated 97 kDa H chain or the parent 145 kDa NT. Because the type A NT is isolated from *C. botulinum* (strain Hall) that is proteo-lytic [8], we extended our study of the time dependent appearance of ~28 and ~18 kDa bands to the type E NT which is produced by nonproteolytic *C. botulinum* (strain, E-43) [8]. The L chain subunit isolated from type E NT [4] but not the H chain, was also found to yield ~28 kDa and ~18 kDa fragments. These and other experimental observations argue against the probability that the breakdown of the L chain at a precise site is mediated by traces of protease that could be present in the purified NT and suggest that the breakdown is spontaneous (see *Discussion*). Further studies on the struct-ure–function relationship of the NT using isolated L chain may take into account the breakdown of the L chain reported here.

Partial amino acid sequences of the ~28 kDa and ~18 kDa fragments of the L chains revealed extensive amino acid sequence homology i), between the antigenically distinct *botulinum* NT types A and E; and ii), among the two *botulinum* NT serotypes which cause flaccid paralysis, and tetanus NT which blocks release of inhibitor neurotransmitter and causes spastic paralysis in tetanus.

Materials and methods

Botulinum NT serotypes A (strain Hall) and E (strain E-43) were produced and purified in our laboratory [9, 10]. The H and L chain subunits of type A NT were separated on ion exchange columns [4]. Because the type E NT is isolated from the bacte-rial culture as a 147 kDa single chain protein [10] it was nicked with trypsin to generate the 147 kDa dichain form in order to separate its subunit chains [4]. Details of nicking with trypsin, inactivation of trypsin, chromatographic separation of the chains, their purity, and partial N-terminal sequences of the subunits of type A and E have already been described [4]. Soybean trypsin inhibitor, aprotinin and phenylmethane sulfonyl fluoride (PMSF), all purchased from Sigma Chem., were used as protease inhibitors.

The L chains of types A and E NTs following separation from their conjugate H chains on QAE–Sephadex columns were recovered in borate–phosphate buffer pH 8.4, containing 2 M urea and at least 10 mM dithiothreitol (DTT) (see [4] for exact compositions of buffer). The ~50 kDa L chains were dialyzed extensively at 8°C, against 0.02 M Na_2HPO_4–NaH_2PO_4 buffer, pH 6.0 con-taining 5 mM DTT and then precipitated with solid ammonium sulfate (39 g/100 ml, Ultrapure, Schwarz/Mann) and stored at 8°C. At suitable inter-vals the protein samples were recovered by centrifu-gation, dissolved in electrophoresis buffer [11] and analyzed by SDS–PAGE [11]. SDS–PAGE of the protein samples in cylindrical gels has already been described [11]. The following proteins (Pharmacia; m.w. indicated in parentheses) were used to estimate the size of the fragments of the L chain; phosphory-lase b (94 000), bovine serum albumin (67 000), ovalbumin (43 000), carbonic anhydrase (30 000), soybean trypsin inhibitor (20 100), α-lactalbumin (14 400). Migration distance plotted against log molecular weight of the proteins provided the calibration curve [12].

For amino acid sequence the Coomassie blue-stained gels were sliced with a razor blade, and the stained protein was recovered from the gel by elec-troelution [13], as was done before [14]. A gas-phase sequencer (Applied Biosystems Model 470A) was used; criteria of identification of phenylthiohy-dantoin derivatives of the amino acids have been described [14].

Relative concentrations of the L chain and its two fragments were estimated from densitometric tracings of the Coomassie blue stained bands in the cylindrical gels at 560 nm using a Gilford linear transport scanner (model 24105) running at 2 cm/min, with recorder chart speed at 5.01 mm/min. The areas in the traced peaks, measured with a compensating polar planimeter (Keuffel & Esser Co., Morristown, NF), represented relative protein concentrations [11].

Results

Progress of breakdown of the 53 kDa L chain of type A NT into ~28 kDa and ~18 kDa fragments during 2, 7, 22 and 45 days at 8°C in 0.02 M phosphate buffer pH 6.0, 5 mM DTT and 60% ammonium sulfate saturation is shown in Figures 1 and 2 (left panels). Analysis of the 145 kDa dichain type A NT stored over compa-
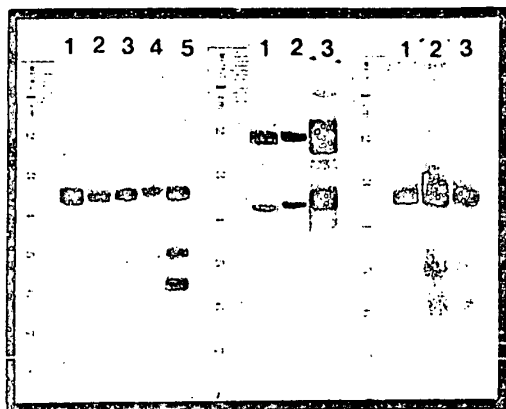

Fig. 1. Photographs of 6 mm cylindrical gels show type A L chain (left panel, gels 1–5), type A NT (middle panel, gels 1–3) and type E L chain (right panel, gels 1–3) and the progress of formation of ~28 and ~18 kDa fragments. The left panel shows appearance of ~28 and ~18 kDa bands at ~50 and ~60 mm, and the type A L chain at 35 mm. The L chain immediately after isolation and after 2, 7, 22, and 45 days are shown in gels 1–5, respectively (densitometric scans of these gels are shown in Fig. 2 left panels a–e, respectively). Direction of migration of bands in gels top to bottom (anode) corresponds to peaks migrating left to right in charts. L chain samples in gels #1–4 were from NT batch #34; gel #5 had L chain isolated from NT batch #31. Loads of protein on gels #1–4 were ~0.02 mg in 400 μl, gel #5 received ~0.06 mg in 150 μl. Because the gels were prepared and run on different days, the locations of the bands do not align perfectly between the gels. Also, the intensity of the bands varies between these gels because they were stained and destained on different days and thus nonidentically. The middle panel shows that type A NT stored for 14 (gel 1), 21 (gel 2) and 42 days (gel 3) dissociated into H and L chains as bands at 18–22 mm and 35–37 mm respectively. Gels 1 and 2 were loaded with ~0.03 mg protein in 100 μl, gel 3 was loaded with twice the amount. Trace impurities are visible at high protein load, but the ~28 and ~18 kDa fragments are not seen. The right panel shows type E L chain immediately after isolation (gel 1) and after 27 (gel 2) and 74 days storage (gel 3). The ~28 kDa and ~18 kDa fragments are at 47–50 mm and 55–60 mm, respectively. Gels 1, 2 and 3 were loaded with ~0.03, ~0.04 and ~0.03 mg protein in 400 μl. NT batches JG-2, 49 and JG-1 were used for gels 1, 2 and 3, respectively. Protein samples were boiled 5 min in the presence of SDS and mercaptoethanol [15]. The 7.5% gels (15 mg acrylamide plus 390 mg *N,N'*-methylenebisacrylamide in 200 ml buffer) were run 5.5 h at 9 mA/gel.

rable period did not yield these two fragments (Figs. 1 and 2, middle panels). The isolated 97 kDa H chain also did not yield these fragments (not shown).

Three separately purified preparations of the type A NT (batch #30, 31 and 34) were used as independent source materials to isolate the L chain. In all cases the pattern of the L chain breakdown was similar. The ~18 kDa fragments generated in these separate preparations in 16 days (batch #34), 30 days (batch #30) and 45 days (batch #31) and the ~28 kDa fragment from one preparation (batch #30, 30 days) were examined for partial amino acid sequences. The freshly prepared L chain, not stored in pH 6.0, was also partially sequenced. The replicate samples of ~18 kDa fragment had identical sequences (Fig. 3).

Fragmentation of the L chain never reached completion in spite of manipulation of the storage conditions in terms of time (up to 45 days for type A and 74 days for type E), temperature (room temperature *vs.* 8°), pH (8.4 *vs.* 6.0), presence or absence of ammonium sulfate at 60% saturation or 2 M urea. Storage of the L chain in the presence of protease inhibitors such as soybean trypsin inhibitor (inhibitor:L chain; 5:100 (w/w), PMSF (2 mM) and aprotinin (each 0.5% w/v) did not prevent or significantly retard progress of fragmentation. The relative proportions of the 53 kDa L chain to the fragments (*i.e.*, ~28 kDa plus ~18 kDa fragments) estimated from 3 separate experiments were as follows: 1:1.93 (batch #31, 45 days old); 1:1.73 (batch #31, 31 days old), and 1:1.27 (batch 34, 16 days old). The ratios of 28 kDa and 18 kDa fragments in these samples were 1:1.61, 1:1.03 and 1:1.63, respectively.

The 50 kDa L chain subunit of type E NT, following its separation from the H chain was also found to break down into ~28 kDa and ~18 kDa fragments (Figs. 1 and 2; right panels). Two separately purified batches of type E (batch #JG-1 and JG-4) were used as independent source material to isolate the L chain and analyze its breakdown. The preparations were stored at 8°C in 2 different ways: i) in 0.02 M phosphate buffer, pH 6.0, in the presence of 5 mM DTT and 60% saturation of ammonium sulfate (39 g/100 ml Ultrapure, Schwartz–Mann) for 74 days; and ii) in borate–phosphate buffer pH 8.4 (see [4] for exact composition) in the presence of 5 mM DTT and 60% saturation of ammonium sulfate for 48 days (urea was removed by dialysis against borate–phosphate
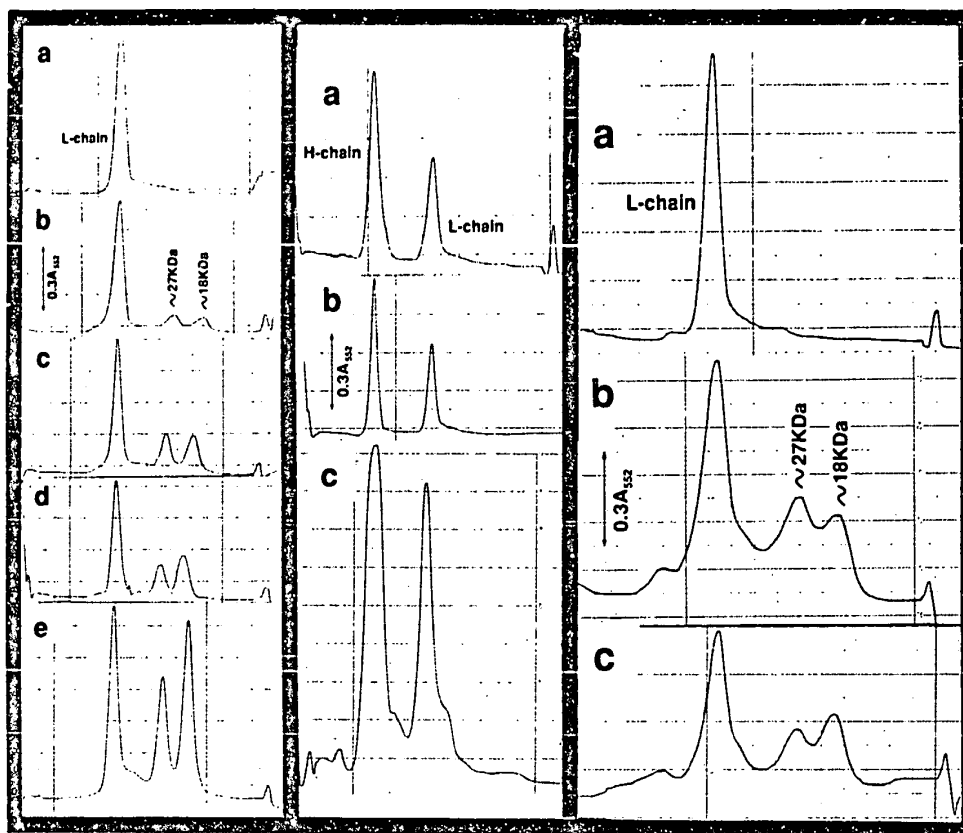
Fig. 2. Densitometric tracings of the cylindrical gels (in Fig. 1) loaded with type A L chain (left panels, a–e), type A NT (middle panels, a–c) and type E L chain (right panels, a–c. Left panels a–e correspond to gels 1–5 in Fig. 1 (left panel). Middle panels a–c are tracings of the gels 1–3, respectively, shown in Fig. 1 (middle panel). Right panels a–c represent gels 1–3, respectively, in Fig. 1 (right panel).

buffer containing DTT). The ~18 kDa fragments from these two preparations and ~28 kDa fragment from one preparation were partially sequenced (Fig. 3). The isolated H chain, its progenitors the dichain NT (nicked with trypsin) and the single chain type E NT (before nicking with trypsin) following long storage under similar conditions did not yield the ~18 and ~28 kDa fragments (results of SDS–PAGE not shown). As mentioned in the case of type A NT, manipulation of storage conditions did not significantly arrest or promote the breakdown of the L chain to completion. At the end of 74 days of storage the relative proportion of the L chain to the fragments (28 kDa plus 18 kDa) from the batch #JG-1 was 1:0.9; and the ratio of 28 kDa:18 kDa fragments was 1:1.45.

Discussion

Examination of the L chain isolated from different preparations of the ~150 kDa NT that were purified from separate batches of bacterial culture showed the following consistent results: i) the ~50 kDa L chain yields ~28 kDa and ~18 kDa fragments; ii) generation of these two fragments is time-dependent but fragmentation does not go to completion; iii) fragmentation results from peptide cleavage (SDS–PAGE and amino acid sequence analyses) at a precise site (the newly released N-terminal on the ~18 kDa fragment is identical in replicate samples); iv) the sites of peptide cleavage on the L chains from types A and E are only 5 amino acid residues apart. Compared to replicate analyses

Neurotoxin	Light chain					
	1	10	20	30	40	49
Botulinum A	P.F.V.N.K.Q.F.N.Y.K.D.P.V.N.G.V.D.I.A.Y.I.K.I.P.N.A.G.Q.M.Q.P.V.K.A.F.K.I.H.N.K.I.W.V.I.P.E.R.D.T.					
Botulinum E	P.K.I.N.S.F.N.Y.N.D.P.V.N.D.R.T.I.L.Y.I.-					
Tetanus	P.I.T.I.N.N.F.R.Y.S.O.P.V.N.N.D.T.I.I.M.M.E.P.P.Y.C.K.G.L.D.I.Y.Y.K.A.F.K.I.I.D.R.I.W.I.V.P.E.R.Y.					

	50	60	70
Botulinum A	F.T.N.P.E.E.G.D.L.N.P.P.P.E.A.K.Q.V.P.X.S.Y.Y.D.-		
Tetanus	E.F.G.T.K.P.E.D.F.N.P.P.S.S.L.I.E.G.A.S.E.Y.Y.D.-		

	~18 kDa fragment				
	1	10	20	30	40
Botulinum A	-Y.E.M.S.G.L.E.V.S.F.E.E.L.R.I.F.G.G.H.D.A.K.F.I.D.S.L.Q.E.N.E.F.R.L.Y.Y.Y.N.K.F.K.D.L.A.S.T.L.-				
Botulinum E	-K.G.I.N.I.E.E.F.L.I.F.G.N.M.D.L.N.I.I.T.V.A.Q.Y.N.D.I.Y.T.N.L.L.N.D.Y.R.K.L.A.X.K.L.-				
Tetanus	-Y.M.Q.H.T.Y.P.I.S.A.E.E.L.F.I.F.G.G.Q.D.A.N.L.I.S.I.D.I.K.N.D.L.Y.E.K.T.L.N.D.Y.K.A.L.A.N.K.L.-				
residues (#259-305)					

Fig. 3. Partial amino acid sequences of the L chains of *botulinum* types A and E and their ~18 kDa fragment compared with tetanus NT. The residues of the light chain of type A and E *botulinum* NT underlined match with the underlined residues in tetanus NT. Tetanus sequence data from [29].

of the ~28 and ~18 kDa fragments from the same preparation of the L chain, our analyses of the L chain isolated from separate batches of the NT allowed a more rigorous examination of the two fragments. Lastly, any conclusion or hypothesis that can be entertained regarding the cause of peptide cleavage is applicable to type A as well as type E.

Why was the fragmentation of the L chain not observed earlier [15, 16, 17]? The L chain subunits of type A and E NTs were first identified in SDS–PAGE [15]. Each dichain NT not reduced with mercaptoethanol migrated as a single band (~150 kDa); following reduction of the disulfide bonds, this band was replaced by two bands (~100 and ~50 kDa). The ~100 and ~50 kDa chains separated electrophoretically in the presence of SDS remained fixed in the gel. Next, for amino acid sequence determinations, the two subunit chains, following recovery from QAE–Sephadex column [4] and precipitation with ammonium sulfate, were processed quickly for Edman degradation [16, 17]. For amino acid analysis [4] the L chain preparations soon after isolation were checked for purity by SDS–PAGE (see Fig. 1 and 3 in [4], dialyzed extensively against water, lyophilized and then acid hydrolyzed. Even if the L chain had fragmented the smaller fragment must have remained inside the dialysis bag, because the sum of the amino acid residues of the lyophilized L and H chains added up to the amino acid content of the 150 kDa dichain NT [4].

We do not have a definitive explanation as to why the fragmentation occurs. Probable causes of the fragmentation, proteolytic or nonproteolytic, are considered here based on experiments, observations and analogies. Although the presence of traces of a protease in the NT preparation is a possibility which is extremely difficult to disprove, the following considerations argue against a protease catalyzed fragmentation: i) during its isolation, the NT was treated with the protease inhibitors aprotinin and PMSF twice, *i.e.* at two different steps of the purification protocol [9]; ii) if a proteolytic activity survived the inhibitors and was copurified with the NT, the protease did not generate the ~28 and ~18 kDa fragments from the dichain NT (composed of L and H chains) or other fragments from the isolated H chain; iii) only the L chain was fragmented, and then only after separation from the H chain. This would be possible if the L chain has a unique amino acid sequence, not present in the H chain, and becomes exposed to the hypothetical protease only after separation from the H chain; iv) the strains of *C. botulinum* that produce types A and E NT are proteolytic and nonproteolytic, respectively [8, 18]. This difference in the physiology between the two strains is also evident from the facts that: type A NT isolated from 96-hr old cultures is a dichain protein (the single chain protein is nicked by an endogenous protease at 1/3 the distance from the N-terminal, see [4, 16, 17]); and the type E NT is recovered as a single

chain molecule. Therefore the type E NT isolated from non-proteolytic culture would be less likely to have the trace protease considered in the case of type A NT. Trypsin, following nicking of the single chain type E to the dichain form, was inactivated with soybean inhibitor and/or PMSF or removed by using soybean trypsin–inhibitor–agarose (Sigma). The dichain type E was the source of the L chain. As mentioned above, a trace of trypsin could survive the inhibitors but then its proteolytic activity must be confined to a unique peptide bond present only when the L chain is isolated, rather than combined with the H chain; v) fragmentation of the L chain from type A or E has never reached completion or near completion (see Fig. 2 and ratio of peaks) in spite of manipulation of solvent, temperature, pH and time.

An explanation alternative to the protease catalyzed fragmentation is that an intrinsic property of the L chain polypeptide causes its breakdown. Studies of various proteins indicate that a simple code in a polypeptide, i.e. certain amino acids and specific sequences, can be the determinant of the polypeptide's degradation. These studies led to the formulation of the N-terminal rule of Varshavasky [19], the PEST hypothesis [20] and the identification of the –KFERQ–sequence [21] and of both aspartyl and asparaginyl peptides as possible hot spots for nonenzymatic breakdown of proteins [22].

The spontaneous degradation at physiological pH and over a range of temperatures was noted in: i) the hexapeptide –Val–Tyr–Pro–Asn–Leu–Ala– (cleavage between Asn and Leu) modeled after sequences of residues #22–27 (–Val–Tyr–Pro–Asn–Gly–Ala–) of adrenocorticotropic hormone [22]; and ii) the αA subunit of bovine α-crystallin (cleavage at residue Asn-101; [23]). The cleavage in both cases occurs following the same pathway, viz. cyclization of Asn to succinimide ([22–24]; and Fig. 1, in [23]). These examples and discussions [22–25] allow us to conjecture that the cause of slow and incomplete breakdown of the L chain is perhaps formation of a succinimide ring from certain Asn or Asp residues, which are plentiful in the NT [4]. The validity of this proposal could be tested using sophisticated and demanding protein chemistry techniques; for example, the truncated L chain, i.e. the ~28 kDa fragment, would be expected to have C-terminal asparagine and also C-terminal aspartic acid amide, because hydrolysis of the labile succinimide takes two paths (see structures VI and VII in Fig. 1 of

[23]). Detection of a C-terminal aspartic acid amide, as Voorter et al. [23]) have suggested, would also eliminate the possibility that a protease cleaves the L chain into ~28 and ~18 kDa fragments.

Why does the L chain break down only after its separation from the H chain? A probable answer is found in Clark's explanation [24]: "asp and asn residues generally exist in native proteins in conformations where the peptide bond nitrogen atom cannot approach the side-chain carbonyl to form a succinimide ring –succinimides may largely form in denatured or partially denaturated proteins in which rotation around the main chain (psi angle) and side chain (chi angle) dihedral angle allow intramolecular imide formation at sites which are conformationally unfavorable in the native structure". Clark [24] also pointed out that "main chain conformations that are conducive to succinimide attack (psi = -120°) are relatively rare in proteins" but in type II' beta hairpin turns this conformation is found for the amino acid residue in the second position. Relevant to this, we note that: i) the L chain, following its separation from the H chain in the presence of 2 M urea and DTT, may not be in a conformation identical to when it is linked to the H chain via disulfide and noncovalent bonds; and ii) secondary structure analysis shows a large content of β-turn in the L chain; 18.75% compared to 13.00% in the H chain and 15.25% in the NT [26].

The structural similarity between botulinum and tetanus NTs, first pointed out in 1977 [27] has became more evident with the availability of their amino acid compositions and partial amino acid sequences ([4] and refs. therein). The L chains of the botulinum and tetanus NTs inhibit exocytosis of catecholamine from chromaffin cells [7, 28]. The complete amino acid sequence of tetanus NT, deduced from the nucleotide sequence, shows that the mature protein (~150,700 Da) and its light (~52,288 Da) and heavy (98,300 Da) chain regions are composed of 1314, 456 and 858 amino acid residues, respectively [29]. Characterizations of the ~28 and ~18 kDa fragments from type A and E botulinum NTs based on partial amino acid sequences are discussed in the context of tetanus NT. The first 20 residues of the ~28 kDa fragment from botulinum type A NT matched exactly with the N-terminal sequence of its L chain, proof that this fragment contains the N-terminal region of the L chain. The sequence of residues 1–73 of the intact L chain includes residues 1–17 identi-

fied earlier [4, 16]; residues 1—46 and residues 1—73 has been described in preliminary reports [30, 31]. The residues underlined in Figure 3 are positionally identical to those of the L chain of tetanus NT. The partial sequence of the ~28 kDa fragment from type E L chain also matched completely with that of the first 20 residues of 50 kDa E L chain [17].

The ~18 kDa fragments from types A and E align with the L chain of tetanus NT beginning at its residue #259 (Tyr) and #264 (Tyr), respectively (Pro, the N-terminus of tetanus NT is counted as #1 residue). This region of the tetanus NT contains the C-terminal half of its L chain (half of 456 residues). The two ~18 KDa fragments can therefore be considered to correspond to the C-terminal half of the L chains of *botulinum* NT types A and E. Homology based on positional identity alone between types A and E is 31%; between *botulinum* type A and tetanus 38.2%; and between *botulinum* type E and tetanus 48.8%. Homology increases if single nucleotide base substitution is considered (not shown here).

Fragmentation of the isolated L chain of *botulinum* NT is not restricted to types A and E. We have observed the same in *botulinum* NT type B, which is currently under investigation.

The mode of action of *botulinum* NT is like those of other dichain toxic proteins (*e.g.* diphteria, ricin) in that its L chain after entering the cell appears to cause the metabolic lesion [7]. Whether the L chain separates from the H chain *in vivo* is not yet known, but the slow *in vitro* fragmentation reported here warrants consideration as in the studies where the isolated L chain was used [5—7].

Acknowledgments

References

1 Syuto B. & Kubo S. (1981) Separation and characterization of heavy and light chains from *Clostridium botulinum* type C toxin and their reconsti-

tution. *J. Biol. Chem.* 256, 3712—3717

2 Kozaki S., Togashi S. & Sakaguchi G. (1981) Separation of *Clostridium botulinum* type A derivative toxin into two fragments. *Jap. J. Med. Sci.* 34, 61—68

3 Kozaki S., Miyazaki S. & Sakaguchi G. (1977) Development of antitoxin with each of two complementary fragments of *Clostridium botulinum* type B derivative toxin. *Infect. Immun.* 18, 761—766

4 Sathyamoorthy V. & DasGupta B.R. (1985) Separation, purification, partial characterization and comparison of the heavy and light chains of *botulinum* neurotoxin types A, B and E. *J. Biol. Chem.* 260, 10461—10466

5 Bandyopadhyay S., Clark A.W., DasGupta B.R. & Sathyamoorthy V. (1987) Role of the heavy and light chains of *botulinum* neurotoxin in neuromuscular paralysis. *J. Biol. Chem.* 262, 2660—2663

6 Maisey E.A., Wadsworth J.D.F., Poulain B., Shone C.S., Melling J., Gibbs P., Tauc L. & Dolly J.O. (1988) Involvement of the constituent chains of *botulinum* neurotoxins A and B in the blockade of neurotransmitter release. *Eur. J. Biochem.* 177, 683—691

7 Bittner M., Holtz R. & DasGupta B.R. (1988) Effects of clostridial neurotoxins on catecholamine release from digitonin-permeabilized chromaffin cells. *Neurosci. Abstr.* 14, Part 2, 68

8 Smith L.Ds. & Sugiyama H. (1988) *in: Botulism* C.C. Thomas, Springfield, IL, 2nd edn.

9 DasGupta B.R. & Sathyamoorthy V. (1984) Purification and amino acid composition of type A *botulinum* neurotoxin. *Toxicon* 22, 415—424

10 DasGupta B.R. & Rasmussen S. (1983) Purification and amino acid composition of type E *botulinum* neurotoxin. *Toxicon* 21, 535—545

11 DasGupta B.R. & Sugiyama H. (1976) Molecular forms of neurotoxins in proteolytic *Clostridium botulinum* type B cultures. *Infect. Immun.* 14, 680—686

12 Weber K. & Osborn M. (1969) The reliability of molecular weight determination by dodecyl sulfate polyacrylamide gel electrophoresis. *J. Biol. Chem.* 244, 4406—4412

13 Hunkapiller M.W., Lujan E., Ostrander F. & Hood L.E. (1983) Isolation of microgram quantities and proteins from polyacrylamide gels for amino acid sequence analysis. *Methods Enzymol.* 91, 227—236

14 DasGupta B.R. & Datta A. (1988) *Botulinum* neurotoxin type B (strain 657): partial sequence and similarity with tetanus toxin. *Biochimie* 70, 811—817

15 DasGupta B.R. & Sugiyama H. (1972) A Common subunit structure in *Clostridium botulinum* type A, B and E toxins. *Biochem. Biophys. Res. Commun.* 48, 108—112

16 Schmidt J.J., Sathyamoorthy V. & DasGupta B.R. (1984) Partial amino acid sequence of the

heavy and light chains of *botulinum* neurotoxin type A. *Biochem. Biophys. Res. Commun.* 119, 900–904

17 Sathyamoorthy V. & DasGupta B.R. (1985) Partial amino acid sequences of the heavy and light chains of *botulinum* neurotoxin type E. *Biochem. Biophys. Res. Commun.* 127, 768–772

18 Das Gupta B.R. & Sugiyama H. (1972) Role of a protease in natural activation of *Clostridium botulinum* neurotoxin. *Infect. Immun.* 6, 587–590

19 Bachmair A., Finley D. & Varshavasky A. (1986) In vivo half-life of a protein as a function of its amino-terminal residue. *Science* 234, 179–186

20 Rogers S., Wells R. & Rechsteiner M. (1986) Amino acid sequences common to rapidly degraded proteins: the PEST hypothesis. *Science* 234, 364–368

21 Dice J.F. (1987) Molecular determinants of protein half-lives in eukaryotic cells. *FASEB J.* 1, 349–357

22 Geiger T. & Clarke S. (1987) Deamidation, isomerization, and racemization at asparaginyl and aspartyl residues in peptides. *J. Biol. Chem.* 262, 785–794

23 Voorter C.E.M., deHaard-Hoekman W.A., van den Oetelaar P.J.M., Bloemendal H. & deJong W.W. (1988) Spontaneous peptide bond cleavage in aging α-crystallin through a succinimide intermediate. *J. Biol. Chem.* 263, 19020–19023

24 Clarke S. (1987) Propensity for spontaneous suc-cinimide formation from aspartyl and asparaginyl residues in cellular proteins. *Int. J. Peptide Protein Res.* 30, 808–821

25 Kossiakoff A.A. (1988) Tertiary structure is a principal determinant to protein deamidation. *Science* 240, 191–194

26 Singh B.R. & DasGupta B.R. (1989) Structure of heavy and light chain subunits of type A *botulinum* neurotoxin analyzed by circular dichroism and fluorescence measurements. *Mol. Cell. Biochem.* 85, 67–73

27 DasGupta B.R. & Sugiyama H. (1977) Biochemistry and pharmacology of *botulinum* and tetanus neurotoxins. *in: Perspectives in Toxinology* (A.W. Bernheimer, ed.), John Wiley & Sons, New York, pp. 87–119

28 Ahnert-Hilger G., Weller U., Dauzeuroth M.E., Habermann E. and Gratzl M. (1989) The tetanus toxin light chain inhibits exocytosis. *FEBS Lett.* 242, 245–248

29 Eisel U., Jarausch W., Gioretzki K., Henschen A., Engels J., Weller U., Hudel M., Habermann E. & Nieman H. (1986) Tetanus toxin: primary structure, expression in *E. coli* and homology with *botulinum* toxins. *EMBO J.* 5, 2495–2502

30 DasGupta B.R., Foley J. & Niece R. (1987) Partial sequence of the light chain of *botulinum* neurotoxin type A. *Biochemistry* 23, 4162

31 DasGupta B.R., Foley J. & Wadsworth C. (1988) *Botulinum* neurotoxin type A: partial sequence of L chain and its two fragments. *FASEB J.* 2, A1750

6   ORI.A47A35

BIOCHIMIE, (1989 Nov-Dec) 71 (11-12) 1193-200.

BIOCHEMICAL JOURNAL, (1990 May 15) 268 (1) 123-8.

INFECTION AND IMMUNITY, (1992 Feb) 60 (2) 518-22.

Sharon L. Turner, Ph.D.
USPTO CM1-10B09
Biotechnology GAU 1647
(703) 308-0056

Characterization of the Neurotoxin Isolated from a *Clostridium baratii* Strain Implicated in Infant Botulism

JUAN A. GIMÉNEZ, MARCELA A. GIMÉNEZ, AND BIBHUTI R. DASGUPTA*

*Department of Food Microbiology and Toxicology, University of Wisconsin,
1925 Willow Drive, Madison, Wisconsin 53706-1187*

Botulism is widely known to result from ingestion of food containing botulinum neurotoxin produced in situ by certain strains of *Clostridium botulinum*. Infant botulism caused by *C. botulinum*, unlike the food-borne intoxication, is the toxicoinfectious form of botulism (S. S. Arnon, p. 331-345, *in* G. E. Lewis, ed., *Biomedical Aspects of Botulism*, 1981). The strain of *Clostridium baratii* implicated in infant botulism produced a neurotoxin that was neutralized with antiserum for botulinum neurotoxin serotype F (J. D. Hall, L. M. McCroskey, B. J. Pincomb, and C. L. Hatheway, J. Clin. Microbiol. 21:654-655, 1985). We developed a procedure to culture the toxigenic *C. baratii* (strain 6341) in dialysis bags and a simple purification scheme (precipitation of 900-ml culture supernatant with ammonium sulfate and two anion-exchange chromatographic steps at pH 5.5 and 8.0) that yielded up to 150 µg of purified neurotoxin. It is an ~140-kDa single-chain protein and has the following sequence of amino acid residues at the N terminus: Pro-Val-Asn-Ile-Asn-Asn-Phe-Asn-Tyr-Asn-Asp-Pro-Ile-Asn-Asn-Thr-Thr-Ile-Leu. Comparison of this amino acid sequence with those of the botulinum neurotoxin serotypes A, B, and E showed 40 to 50% identical residues in comparable positions. The specific toxicity of the neurotoxin, $\sim 2 \times 10^6$ 50% lethal doses for mice per mg of protein injected, was not enhanced significantly by mild trypsinization, although the protease cleaved the neurotoxin within a disulfide loop that generated at least two primary fragments, ~47 and ~86 kDa, that remained linked by an interchain disulfide. These two fragments resembled the light and heavy chains of the well-characterized neurotoxin serotypes A, B, C, D, E, and F produced by *C. botulinum*.

Clostridium botulinum and *C. tetani* have been well recognized, until recently, as the only two bacterial species that produce the two extremely poisonous proteins—botulinum and tetanus neurotoxins (18). The former causes flaccid paralysis in botulism, and the latter causes spastic paralysis in tetanus (16). Botulinum neurotoxin is found in nature as seven antigenically distinguishable proteins (serotypes A, B, C_1, D, E, F, and G); tetanus neurotoxin occurs as a single serotype. The neurotoxins of these two *Clostridium* species have strikingly similar primary structures, structure-function relationships, and modes of action; however, their primary sites of action are different. Botulinum neurotoxin acts at neuromuscular junctions, and tetanus neurotoxin acts at inhibitory synapses in the central nervous system (for reviews of the pharmacology of these two proteins, see references 16 and 27).

Certain strains of *C. baratii* and *C. butyricum* responsible for human infant botulism (2) cases were isolated around 1985. The neurotoxins produced in cultures of these clostridial species were neutralized by anti-botulinum neurotoxin sera (17, 24); the genetic identities of the *C. baratii* and *C. butyricum* strains were confirmed (32). These findings have profound implications, one of which is that non-*C. botulinum* species found in the normal flora of infant intestines may acquire and express the neurotoxin gene. This elicited for us an immediate question: how do the neurotoxins produced by *C. baratii* and *C. butyricum* compare structurally with those produced by *C. botulinum*? The neurotoxin elaborated by *C. butyricum* was purified (15). and its partial amino acid sequence (13) and molecular topography (28) were determined; these and a few other features were

compared with those of botulinum neurotoxin type E (13, 15, 21, 28). We now report partial characterization of the neurotoxin produced by *C. baratii* 6341 (to be referred to henceforth as baratii neurotoxin) in terms of (i) molecular size, (ii) partial amino acid sequence, (iii) toxicity (50% lethal doses [LD_{50}] for mice per mg of protein injected), and (iv) enhancement of toxicity (activation) after mild trypsinization. We also compare these features with those of the botulinum neurotoxins.

MATERIALS AND METHODS

Culture media. Four different media of the following compositions were used to maintain the stock culture and to produce the neurotoxin: medium A, 12.5% solid meat media (Difco), 0.3% glucose, 0.2% soluble starch (Difco), and 0.05% sodium-thioglycolate (Sigma); medium B, 2% Trypticase-peptone (BBL, Becton Dickinson Microbiology Systems), 1% glucose, 0.5% yeast extract (BBL), 0.2% soluble starch, and 0.05% sodium-thioglycolate; medium C, 1% glucose and 0.05% sodium-thioglycolate; medium D, 4% Trypticase-peptone, 1% yeast extract, and 0.05% sodium-thioglycolate. In each case all the ingredients were dissolved (wt/vol) in 0.1 M sodium potassium phosphate buffer, pH 7.0.

Bacterial culture. One milliliter of *C. baratii* (strain 6341), grown in cooked-meat medium and kindly provided by Charles Hatheway (Centers for Disease Control, Atlanta, Ga.), was inoculated into 250 ml of medium A in a 500-ml flask and incubated for 5 days at 37°C. The bacterial cells in the culture fluid were packed into a small volume by centrifugation (10,000 × *g*, 20 min). suspended in sterile 20% skim milk (Difco; dissolved in water), and then distributed in 2-ml
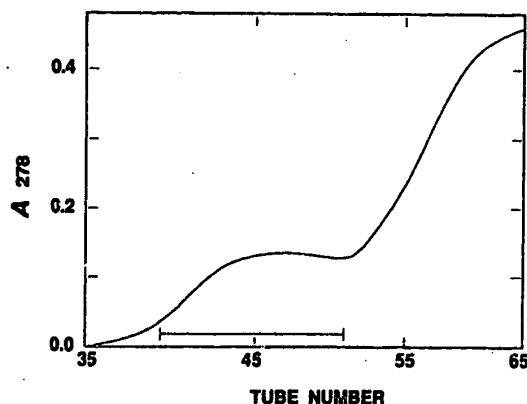
* Corresponding author.

FIG. 1. Chromatography of culture supernatant concentrate on a DEAE-Sephadex A-50 column (2.5 by 60 cm) equilibrated and eluted with 50 mM sodium citrate buffer. pH 5.5. Chromatography was done at room temperature. The flow rate was 25 ml/h, and the fraction volume was 2.4 ml. The fractions containing the first peak were pooled and are indicated with a bar.
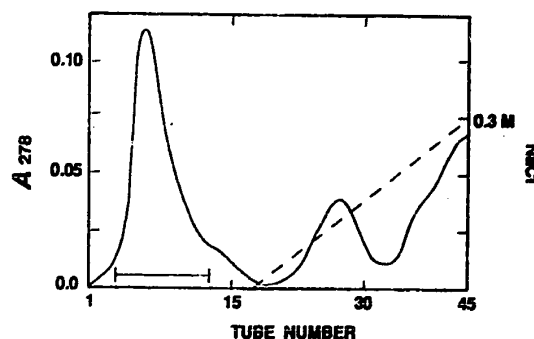

FIG. 2. Chromatography on a DEAE-Sephadex A-50 column (1.8 by 8 cm) equilibrated and eluted at 4°C with 20 mM sodium phosphate buffer (pH 8.0). Flow rate was 12 ml/h, and the fractions were of 2 ml. The salt gradient, indicated with a dashed line, was made with 50 ml of 20 mM sodium phosphate buffer, pH 8.0, and 50 ml of the same buffer but containing 0.3 M NaCl. The neurotoxin eluted isocratically and was recovered in the first peak (neurotoxin pool is indicated with a bar).

vials as stock culture. These were stored at −70°C or lyophilized and stored at 4°C.

The neurotoxin was produced in dialysis bags (31) with 32- or 100-mm-flat-width and 100-cm-long dialysis tubings (Spectrapor; D1615-2, 12- to 14-kDa cutoff, or 132670, 6- to 8-kDa cutoff limit). Approximately 30 ml of medium B in a 50-ml screw-cap glass tube was inoculated with 0.2 ml of stock culture (thawed or rehydrated with medium B). After ~12 h of incubation at 37°C, 1 ml of this culture was transferred to 30 ml of medium B in a glass tube and further incubated overnight at 37°C. One milliliter of this culture was inoculated into 150 ml of medium C contained in a dialysis bag which was immersed in a 1,000-ml glass cylinder containing 500 ml of medium D. After 5 days of incubation at 37°C, the contents of dialysis tubes from six parallel cylinders were pooled and centrifuged (10,000 × g, 60 min). The supernatant was 60% saturated with ammonium sulfate (390 mg/ml) and kept at 4°C overnight. Our one attempt to harvest more neurotoxin from a larger volume of culture fluid by using the larger-diameter dialysis bag (100 versus 32 mm) of identical length (100 cm) gave lower yield. Porosity of the bags (i.e., molecular size cutoff and volumes inside and outside of the bags were different; whether these were factors in the neurotoxin yield remains to be explored.

Neurotoxin purification. The material precipitated with ammonium sulfate was recovered by centrifugation (10,000 × g, 45 min); dissolved in 15 ml of 50 mM citric acid-disodium citrate buffer, pH 5.5; and then dialyzed against this buffer for several hours at 4°C. The dialyzed solution (~25 ml) was centrifuged (15,000 × g, 60 min at 25°C); the supernatant was loaded into a DEAE-Sephadex A-50 column (2.5 by 60 cm), equilibrated, and eluted with the 50 mM citrate buffer, pH 5.5. The UV absorbing material (indicated with a bar in Fig. 1) eluting as a wide peak ahead of a large peak was pooled; dialyzed against 30 mM sodium phosphate buffer, pH 7.0, for 5 h at 4°C; and finally precipitated with ammonium sulfate (390 mg/ml). After 16 h at 4°C, the precipitate was recovered by centrifugation (10,000 × g, 30 min, 4°C); dissolved in 5 ml of 20 mM sodium phosphate buffer, pH 8.0; and dialyzed against this buffer for 16 h at 4°C. Turbidity that appeared during dialysis was removed by

centrifugation. The clear supernatant was loaded onto a DEAE-Sephadex A-50 column (0.9 by 6 cm) equilibrated at 4°C with the pH 8.0 buffer. The UV absorbing material that did not bind to the column and emerged as the first peak (Fig. 2) was pooled, precipitated with ammonium sulfate (390 mg/ml), and stored at 4°C.

SDS-PAGE. Protein samples were analyzed by sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE; 10% acrylamide, 0.89% cross-linking) in a discontinuous buffer (22) and with the stain Coomassie blue R-250 (Sigma). Their molecular masses were estimated from the relative electrophoretic migration of marker proteins (phosphorylase b, bovine serum albumin, ovalbumin, and carbonic anhydrase, which were 94, 67, 43, and 30 kDa, respectively [Pharmacia]) as well as the well-characterized botulinum neurotoxin serotypes A, B, and E (5).

Electroblotting and amino acid sequence determinati... The baratii neurotoxin after electrophoresis on 6% linear acrylamide gel with SDS (1.6-mm-thick, 10- by 10-cm mini-islab; 40 mA for 90 min) by using the discontinuous buffer system (22) was electrotransferred (0.5 A, 60 min) to a polyvinylidene difluoride (Millipore, Bedford, Mass.) membrane by the method of Matsudaira (23) with the minor modifications reported before (12). The protein blotted onto the polyvinylidene difluoride membrane was stained with Coomassie blue, air dried, excised out with a razor blade, and stored at −20°C. For amino acid sequencing, Applied Biosystems Instrument pulse-liquid phase (models 475A and 477A) protein sequencers were used. Phenylthiohydantoin amino acids were identified with an on-line phenylthiohydantoin amino acid analyzer and a data analyzer system (Applied Biosystems model 120-A and 900-A, respectively).

Hemagglutination assay. Human and rabbit erythrocytes, recovered with citrate as an anticoagulant, were washed with 75 mM sodium phosphate buffer, pH 7.2, containing 75 mM NaCl. Microtiter plate (Dispo V plate; Scientific Products) wells were loaded with 50 μl of 0.5% blood cell suspension and 50 μl of test proteins serially diluted in the pH 7.2 buffer. The hemagglutinin protein chromatographically separated from botulinum type A neurotoxin (30) was used as positive control. The plates were read for hemagglutination after 8 h at 6°C.

Toxicity and neutralization with antisera. T xicity was
determined by intraven us injecti n f 0.1 ml of s mple into
the tail vein f white mice (~22 g); time (in minutes) until
death of mice after injection is a functi n of the toxic
potency (4). Ability of the antisera prepared against botuli-
num neurotoxin serotype A, B, E, or F to neutralize baratii
neurotoxin was examined as follows: 50 μl of the neurotoxin
(containing 25 LD₅₀) was incubated with 50 μl of anti-type A.
B, E, or F serum (enough to neutralize 1,000 LD_{50} of the
homologous neurotoxin) for 1 h at 37°C and then injected
intravenc ily or intraperitoneally in mice. Anti-type A. B,
and E neurotoxin sera were produced in rabbits as described
before (8); anti-type F serum, produced in burros, was a gift
from C. Hatheway.

Effects of trypsinization on covalent structure and toxicity.
Baratii neurotoxin (~0.1 mg/ml) was incubated with trypsin
(Sigma; type XIII, lot 67F-8045; ratios [wt/wt], 50:1, 25:1,
10:1, and 2:1) at pH 6.0 or 7.0 (0.1 M sodium-phosphate) or
7.5 or 8.0 (0.1 M Tris-hydrochloride) for 5 to 60 min at 30 to
35°C. Tryptic digestions were stopped with soybean trypsin
inhibitor (Sigma; type I-S, lot 85C-8057); enzyme-to-inhibi-
tor ratio was 1:3 (wt/wt). The digests and the control
(neurotoxin incubated without trypsin) were diluted with 50
mM sodium phosphate buffer, pH 6.0, containing 0.2%
gelatin (Difco) and then tested for toxicity. Trypsinized
neurotoxins with and without dithiothreitol (DTT) reduction
were analyzed by SDS-PAGE.

Protein concentration was determined with bicinchoninic
acid (29) (by using the protein determination kit from Pierce,
Rockford, Ill.). Bovine serum albumin was used as the
reference protein.

RESULTS AND DISCUSSION

C. baratii 6341, cultured conventionally (i.e., without
dialysis bag), produced amounts of the neurotoxin (<500
LD_{50}/ml) that were too low for biochemical studies. Dialysis
bag culturing (31) improved the yield to >5,000 LD_{50}/ml,
which was lower than the range from 10⁵ to 10⁸ LD_{50}/ml
found in cultures of *C. botulinum* types A, E, and F (9, 14,
33). The baratii neurotoxin was purified (2 × 10⁶ LD_{50}/mg)
more than 1,000 times from the culture supernatant (1.5 ×
10³ LD_{50}/mg). The average yield of ~0.15 mg of purified
baratii neurotoxin recovered from ~900-ml culture com-
pared with ~1.0, ~1.5, and 0.63 mg of pure botulinum
neurotoxin types A, E, and F per liter, respectively, from
conventional culture (9, 14, 33) indicates low toxigenicity of
C. baratii 6341 rather than poor recovery of the protein
through purification steps.

The SDS-PAGE analysis shows progress through the
purification steps (Fig. 3). The pool from the pH 5.5 column
(Fig. 1) had 140-, 120-, and 60-kDa protein bands along with
two or three smaller-size proteins (Fig. 3, lane 3). Chroma-
tography at pH 8.0 (Fig. 2) removed the 120-kDa and most of
the other proteins from the 140-kDa neurotoxin. The baratii
neurotoxin did not bind to a DEAE-Sephadex A-50 column
equilibrated with 20 mM sodium phosphate buffer, pH 8.0.
We tested QAE-Sephadex A-50 and Mono-Q resin (Fast
Protein Liquid Chromatography; Pharmacia) equilibrated
with the same buffer. Both were unsatisfactory. The neuro-
toxin preparation partially bound to the QAE gel and eluted
as a broad peak under a gradient of increasing NaCl concen-
tration (not shown). The Mono-Q column bound the neuro-
toxin; an increasing salt gradient eluted a single peak con-
taining a significant amount of impurities (not shown).
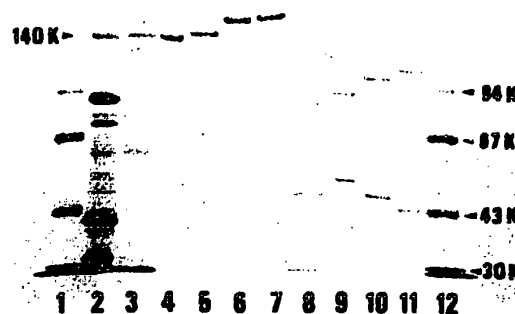
FIG. 3. PAGE pattern of toxin samples. Lanes: 1 and 12, molec-
ular mass markers: 2, crude baratii neurotoxin before chromatogra-
phy at pH 5.5 column: 3, neurotoxin purified through DEAE-
Sephadex at pH 5.5: 4, neurotoxin from the pH 8 chromatography:
5, 6, and 7, botulinum type A. B, and E neurotoxins, respectively
(the two extra bands below type A neurotoxin in lane 5 are
impurities): 8, 10, and 11, baratii type B and E neurotoxins (same as
in lanes 4, 6 and 7, respectively) after trypsinization and reduction
with DTT: 9, botulinum type A neurotoxin reduced with DTT.

DEAE-Sephadex was therefoe chosen for subsequent
work.

The neurotoxin obtained from the DEAE-Sephadex col-
umn (Fig. 3, lane 4) did not change its electrophoretic
migration after reduction with DTT. Thus, the 140-kDa
neurotoxin appears to be a single-chain protein and not
composed of two polypeptides linked by a disulfide bond(s).
Determination of amino acid sequence of the neurotoxin
prepared from two independent batches yielded Pro as the
only N terminus (Table 1). The first preparation was ana-
lyzed for the first 32 residues, and the second batch was
analyzed for the first 19 residues; identical results were
obtained.

The 120-kDa protein (the band migrating immediately
ahead of the 140-kDa neurotoxin in lanes 2 and 3 of Fig. 3)
found associated with the neurotoxin after chromatography
at pH 5.5 was removed after chromatography at pH 8.0 (Fig.
3, lane 4). This protein appears analogous to the nonneuro-
toxic proteins possessing high. low, or no hemagglutinating
activity and found tightly associated with all botulinum
neurotoxin serotypes (25, 30). For example, the nonneuro-
toxic proteins associated with type A and F neurotoxins
have high and low hemagglutinating activity, respectively:
the 134-kDa protein associated with type E has no hemag-
glutinating activity (30). These nonneurotoxic proteins have
other common characteristic features. They remain associ-
ated with the neurotoxins at acidic pH but can be separated
at alkaline pH and high ionic strength, and also the nonneu-
rotoxic proteins elute from an anion-exchange column at an
ionic strength that is higher than that which elutes the
neurotoxin. Baratii neurotoxin was freed of the 120-kDa
protein with 20 mM phosphate buffer, pH 8.0. The 120-kDa
protein was eluted with an increasing salt gradient (Fig. 2).
No hemagglutinating activity was found in the bacterial
culture. concentrated protein samples, or chromatographic
fractions. Absence of hemagglutinating activity resembles

TABLE 1. N-terminal amino acid residues of various *Clostridium* neurotoxins

Neurotoxin	Reference	Amino acid residue			% Homology
		10	20	30	
Baratii[a]		P V N I N N F N Y N D P I N N T T I L Y M K M P Y Y Y D S N K Y			
Botulinum					
Type A	3	P F V N K Q F N Y K D P V N G V D I A Y I K I P N A G Q M Q P V			34
Type B	5	P V T I N N F N Y N D P I D N N N I I M M E P P F A R G M G R Y			53
Type E	5	P - K I N S F N Y N D P V N D R T I L Y I K P G G C Q E F Y K S			50
Butyricum	13	P - T I N S F N Y N D P V N N R T I L Y I K P G ü X Q Q F Y K S			53
Tetanus	11	P I T I N N F R Y S D P V N N D T J I M M E P P Y C K G L D I Y			50

[a] The amino acid sequenator log is as follows: the theoretical initial yields (in picomoles) and combined amino acid repetitive yield (in percent) of the first batch (32 residues identified; LPS 1351) were 58 and 92%. respectively. and for the second batch (19 residues identified: MLK 5-7-90) were 10 and 91%. respectively.

the properties of products of type E culture. The relative proportions of baratii neurotoxin and the 120-kDa protein did not appear constant in the different preparations: the neurotoxin was always in a larger amount.

Specific toxicity of the purified baratii neurotoxin was ~2 × 10^6 LD$_{50}$/mg. Anti-botulinum type A or B serum did not neutralize toxicity. Complete neutralization of toxicity was attained with anti-type F serum. as expected (17). We also found that baratii neurotoxin is partially neutralized (delayed death in mouse lethality assay) with anti-type E serum (Hatheway also noted this [18a]).

Toxicity of the baratii neurotoxin was not significantly enhanced by trypsin digestion. Limited digestion of the neurotoxin with trypsin (20:1 [wt/wt]. 30 min at 35°C) followed by reduction of disulfide bonds with DTT generated two major fragments (86 and 47 kDa) and at least three other minor fragments (<140. ~60. and ~43 kDa) as shown in Fig. 3 (lanes 4 and 8). Higher trypsin concentration or longer digestion produced further cleavage of the 140-kDa neurotoxin as well as the 86- and 47-kDa fragments.

Generation of 86- and 47-kDa fragments from the 140-kDa protein by trypsinization and separation of the two fragments after reduction with DTT suggest that trypsin cleaved the neurotoxin at about one-third the distance from its N or C terminus and within a disulfide loop. We infer that 86- and 47-kDa fragments represent the heavy and light chains. respectively. of the neurotoxin. This inferred structure is consistent with the structures of different serotypes of botulinum neurotoxin as well as tetanus neurotoxin. Each of these neurotoxins is synthesized as an ~150-kDa single-chain protein. Proteolytic cleavage (nicking) converts these proteins to ~150-kDa dichain proteins. each composed of an ~100-kDa heavy chain and an ~50-kDa light chain that remain linked by noncovalent bonds and a disulfide bond. and the light chain retains the N-terminal segment of the parent single chain (5). Proteases endogenous to the bacteria or exogenous. such as trypsin and endoproteinase Lys-C. nick the single-chain neurotoxin to the dichain form (20). Botulinum neurotoxin serotypes A. B. and E isolated from 96-h-old bacterial cultures are. respectively. nicked (dichain). partially nicked (mixture of 80 to 90% single and 10 to 20% dichain). and unnicked (single-chain) proteins (5). In Fig. 3 the 140-kDa baratii neurotoxin (lane 4) is compared with ~150-kDa type A. B. and E neurotoxins (lanes 5. 6. and 7. respectively); lanes 9. 10. and 11 exhibit the heavy and light chains of types A. B. and E. respectively. (Type B and E neurotoxins were trypsinized to generate the dichain forms [26].) In lane 8 the 140-kDa band represents the residual amount of undigested baratii neurotoxin. Our limited attempts to generate a homogeneous population of

dichain protein by confining cleavage of the single-chain neurotoxin at only one peptide bond were not successful: additional cleavages occurred. The thin band immediately below the 140-kDa band is probably a fraction of the neurotoxin population that was cleaved near N or C termini but was not nicked. One band located between the heavy and light chain and the one below the light chain are products of cleavages at sites other than nicking. Similar additional fragmentations. in addition to nicking. have been observed with type B and E neurotoxins treated with trypsin and endoproteinase Lys-C (20. 26).

A short stretch at one-third the distance from the N terminus of botulinum type A and E and tetanus neurotoxins is rich in Arg and Lys residues (1. 7. 12). Trypsin or the protease endogenous to the neurotoxin-producing bacteria cuts multiple bonds within a range of few residues. Homogeneity of the N terminus of the heavy chain therefore depends on precise conditions of proteolysis and the end point of proteolysis (1. 7. 12). The 86-kDa heavy and 47-kDa light chains of baratii neurotoxin produced by trypsinization (Fig. 3. lane 8) were heterogenous (SDS-PAGE showed more than a single band): therefore. we did not determine their N-terminal sequences.

The 86-kDa fragment generated by trypsinization (Fig. 3) is. we believe. the heavy chain segment of the neurotoxin. After verifications of its identity. this material could provide interesting opportunities in the structure-function relationship studies. Electrophoretic migration (Fig. 3) indicates that baratii neurotoxin heavy chain is smaller in molecular mass than those of types A. B. and E. while the baratii neurotoxin light chain is smaller than that of type A. similar to that of type B. and larger than that of type E. The heavy chain appears to have at least three roles in the neurotoxin's pathway of intoxication. (i) In the 150-kDa dichain neurotoxin the ~50-kDa light chain remains linked to the N-terminal half of the heavy chain: unless the heavy chain delivers it into the cytosol. the light chain cannot poison the neurotransmitter secretory cells. (ii) The C-terminal half of the heavy chain is thought to have the binding site for receptors on the target cell membrane (6. 19). (iii) After receptor-mediated endocytosis. the neurotoxin. inside the cell. initially remains encapsulated by the endosomes. The N-terminal half of the heavy chain forms channels in the endosomal membrane through which the light chain exits into the cytosol (see reference 6 and references therein). The baratii neurotoxin heavy chain is smaller than the heavy chains of type A. B. and E neurotoxins: therefore. its structure-function analysis would be highly instructive in defining which precise segments of the heavy chain play what role.

The baratii neurotoxin purified from 4-day-old culture was

in the single-chain form; this feature differs from the dichain type F neurot xin is lated from 4- t 5-day-old culture of proteolytic strains f C. botulinum type F (10, 33, 34).

The mechanisms by which C. baratii (strain 6341) acquires the neur toxin gene and expresses the pr tein are unknown. The amino acid sequence reported here can be exploited to construct DNA probes which could be valuable in answering some of the genetic questions regarding the location of the neurotoxin gene and its exchange or acquisition by the nonneurotoxic C. baratii and toxigenic C. botulinum.

ACKNOWLEDGMENTS

REFERENCES

1. Ahnert-Hilger, G., M. E. Dauzenroth, E. Habermann, A. Henschen, K. Kriegistein, F. Mauler, and U. Weller. 1990. Chains and fragments of tetanus toxin, and their contribution to toxicity. J. Physiol. 84:229–236.

2. Arnon, S. S. 1981. Infant botulism: pathogenesis, clinical aspects and relation to crib death, p. 331–345. In G. E. Lewis (ed.), Biomedical aspects of botulism. Academic Press, Inc., San Diego.

3. Binz, T., H. Kurazono, M. Wille, J. Frevert, K. Wernars, and H. Niemann. 1990. The complete sequence of botulinum neurotoxin type A and comparison with other clostridial neurotoxins. J. Biol. Chem. 265:9153–9158.

4. Boroff, D. A., and U. Fleck. 1966. Statistical analysis of rapid in vivo method for the titration of the toxin of Clostridium botulinum. J. Bacteriol. 92:1580–1581.

5. DasGupta, B. R. 1989. The structure of botulinum neurotoxin, p. 53–67. In L. L. Simpson (ed.), Botulinum neurotoxin and tetanus toxin. Academic Press, Inc., San Diego.

6. DasGupta, B. R. 1990. Structure and biological activity of botulinum neurotoxin. J. Physiol. 84:220–228.

7. DasGupta, B. R., and M. L. Dekleva. 1990. Botulinum neurotoxin type A: sequence of amino acids at the N-terminus and around the nicking site. Biochimie 72:661–664.

8. DasGupta, B. R., and S. Rasmussen. 1984. Effect of diethylpyrocarbonate on the biological activities of botulinum neurotoxin types A and E. Arch. Biochem. Biophys. 232:172–178.

9. DasGupta, B. R., and V. Sathyamoorthy. 1984. Purification and amino acid composition of type A botulinum neurotoxin. Toxicon 22:415–424.

10. DasGupta, B. R., and H. Sugiyama. 1977. Single chain and dichain forms of neurotoxin in type F Clostridium botulinum culture. Toxicon 15:466–471.

11. Eisel, U., W. Jarausch, K. Goretzki, A. Henschen, J. Engels, U. Weller, M. Hudel, E. Habermann, and H. Nieman. 1986. Tetanus toxin: primary structure, expression in E. coli and homology with botulinum toxins. EMBO J. 5:2495–2502.

12. Giménez, J. A., and B. R. DasGupta. 1990. Botulinum neurotoxin type E fragmented with endoproteinase Lys-C reveals the site trypsin nicks and homology with tetanus neurotoxin. Biochimie 72:213–217.

13. Giménez, J. A., J. Foley, and B. R. DasGupta. 1988. Neurotoxin type E from Clostridium botulinum and C. butyricum: partial sequence and comparison, abstr. 8447. FASEB J. 2:A1750.

14. Giménez, J. A., and H. Sugiyama. 1987. Simplified purification method for Clostridium botulinum type E toxin. Appl. Environ. Microbiol. 53:2827–2830.

15. Giménez, J. A., and H. Sugiyama. 1988. Comparison of toxins of

16. Habermann, E., and F. Dreyer. 1986. Clostridial neurotoxins: handling and action at the cellular and molecular level. Curr. Top. Microbiol. Immunol. 129:93–179.

17. Hall, J. D., L. M. McCroskey, B. J. Pincomb, and C. L. Hatheway. 1985. Isolation of an organism resembling Clostridium baratii which produces type F botulinal toxin from an infant with botulism. J. Clin. Microbiol. 21:654–655.

18. Hatheway, C. L. 1990. Toxigenic clostridia. Clin. Microbiol. Rev. 3:66–98.

18a.Hatheway, C. L. Personal communication.

19. Kozaki, S., A. Miki, Y. Kamata, J. Ogasawara, and G. Sakaguchi. 1989. Immunological characterization of papain-induced fragments of Clostridium botulinum type A neurotoxin and interaction of the fragments with brain synaptosomes. Infect. Immun. 57:2634–2639.

20. Kozaki, S., Y. Oga, Y. Kamata, and G. Sakaguchi. 1985. Activation of Clostridium botulinum type B and E derivative toxins with lysine-specific proteases. FEMS Microbiol. Lett. 27:149–154.

21. Kozaki, S., J. Onimaru, Y. Kamata, and G. Sakaguchi. 1991. Immunological characterization of Clostridium butyricum neurotoxin and its trypsin-induced fragment by use of monoclonal antibodies against Clostridium botulinum type E neurotoxin. Infect. Immun. 59:457–459.

22. Laemmli, U. K. 1970. Cleavage of structural proteins during the assembly of the head of bacteriophage T4. Nature (London) 227:680–685.

23. Matsudaira, P. 1987. Sequence from picomole quantities of proteins electroblotted onto polyvinylidene difluoride membranes. J. Biol. Chem. 262:10035–10038.

24. McCroskey, L. M., C. L. Hatheway, L. Femicia, B. Pasolini, and P. Aureli. 1986. Characterization of an organism that produces type E botulinal toxin but which resembles Clostridium butyricum from the feces of an infant with type E botulism. J. Clin. Microbiol. 23:201–202.

25. Sakaguchi, H. 1983. Clostridium botulinum neurotoxins. Pharmacol. Ther. 19:165–194.

26. Sathyamoorthy, V., and B. R. DasGupta. 1985. Separation, purification, partial characterization and comparison of the heavy and light chains of botulinum neurotoxin types A, B, and E. J. Biol. Chem. 260:10461–10466.

27. Simpson, L. L. 1989. Peripheral actions of botulinum toxins, p. 153–178. In L. L. Simpson (ed.), Botulinum neurotoxin and tetanus toxin. Academic Press, Inc., San Diego.

28. Singh, B. R., J. A. Giménez, and B. R. DasGupta. 1991. Comparative molecular topography of botulinum neurotoxins from Clostridium butyricum and Clostridium botulinum type E. Biochim. Biophys. Acta 1077:119–126.

29. Smith, P. K., R. I. Krohn, G. T. Hermanson, A. K. Mallia, F. H. Gartner, M. D. Provenzano, E. K. Fujimoto, N. M. Goeke, B. J. Olson, and D. C. Klenk. 1985. Measurement of protein using bicinchoninic acid. Anal. Biochem. 150:76–85.

30. Somers, E., and B. R. DasGupta. 1991. Clostridium botulinum types A, B, C_1 and E produce proteins with or without hemagglutinating activity: do they share common amino acid sequences and genes? J. Protein Chem. 10:415–425.

31. Sterne, M., and L. M. Wentzel. 1950. A new method for the large-scale production of high-titre botulinum formol-toxoid types C and D. J. Immunol. 65:175–183.

32. Suen, J. C., C. L. Hatheway, A. G. Steigerwalt, and D. J. Brenner. 1988. Genetic confirmation of identities of neurotoxigenic Clostridium baratii and Clostridium butyricum implicated as agents of infant botulism. J. Clin. Microbiol. 26:2191–2192.

33. Wadsworth, J. D. F., M. Desai, H. S. Tranter, H. J. King, P. Hambleton, J. Melling, J. O. Dolly, and C. C. Shone. 1990. Botulinum type F neurotoxin. Large scale purification and characterization of its binding to rat cerebrocortical synaptosomes. Biochem. J. 268:123–128.

34. Yang, K. H., and H. Sugiyama. 1975. Purification and properties of Clostridium botulinum type F toxin. Appl. Microbiol. 29:598–603.

8

QP501.B43

From:	Turner, Sharon
Sent:	Thursday, May 31, 2001 12:43 PM
To:	STIC-ILL
Subject:	08981087

Please provide

J. Clin. Microbiol., 31:2255-62, 1993

FEMS Microbiol., Lett. 108:175-182, 1993

Biochem. Biophys. Res Comm., 183:107-113, 1992

J. Gen. Microbiol., 137:519-525, 1991

Sharon L. Turner, Ph.D.
USPTO CM1-10B09
Biotechnology GAU 1647
(703) 308-0056

1

SEQUENCES OF THE BOTULINAL NEUROTOXIN E DERIVED FROM CLOSTRIDIUM BOTULINUM TYPE E (STRAIN BELUGA) AND CLOSTRIDIUM BUTYRICUM (STRAINS ATCC 43181 AND ATCC 43755)

Sylvie Poulet[1], Daniel Hauser[1], Martin Quanz[2], Heiner Niemann[2], and Michel R. Popoff[1]*

[1]Unité des Antigènes Bactériens, Institut Pasteur, 25 rue du Docteur Roux, 75724 Paris Cedex 15, France

[2]Federal Research Center for Virus Diseases of Animals, W-7400 Tübingen, Federal Republic of Germany

SUMMARY: Recently, it has been shown that two Clostridium butyricum strains (ATCC 43181 and ATCC 43755), isolated from cases of infant botulism, produce a botulinal neurotoxin type E (BoNT/E). Here we have determined the nucleotide sequences of the BoNT/E genes of these two C. butyricum strains and from C. botulinum E strain Beluga. We show that the sequences of the BoNT/E genes from the two C. butyricum strains are identical and differ in only 64 positions resulting in 39 amino acid changes (97% identity at the amino acid level) from that derived from C. botulinum. Our data suggest a transfer of the BoNT/E gene from C. botulinum to the originally nontoxigenic C. butyricum strains.   © 1992 Academic Press, Inc.

The clostridial neurotoxins are highly potent protein toxins that inhibit neurotransmitter release at various synapses. These neurotoxins consist of tetanus toxin (TeTx) and seven serologically distinct botulinal neurotoxins designated BoNT/A, B, C1, D, E, F, and G, all of which are both structurally and functionally closely related. All of them are synthesized as single chain polypeptides of about 150 kDa which are proteolytically activated into di-chain derivatives constituted of a light (L) (Mr ap. 50 000), and a heavy (H) chain (Mr ap. 100 0000) linked by a single disulfide bridge (1).

*To whom correspondence should be addressed.

A three step model has been proposed to explain the development of neurotoxicity (2). This model involves binding, internalization and intraneuronal sorting, and the actual poisoning of the nerve terminal. The heavy chains mediate the binding of toxins to cell receptors, while the light chains appear to enter target cells, and cause some internal changes (3,4). It is not known how the different heavy chains reach and affect different kinds of synapses, nor what is the molecular nature of the change caused by the light chains. However, these questions may be approached by sequence analyses and the identification of conserved domains. The genes encoding TeTx, BoNT/A, C1, and D have been sequenced previously (5, 6, 7, 8, 9). It has been shown that the BoNT/E gene is chromosomally located in *C. botulinum* E strain Beluga, and a partial sequence was established by cloning and sequencing of an EcoRI fragment (7).

Recently, it has been reported that *Clostridium* strains quite different from *C. botulinum* can produce BoNT. Thus, two *Clostridium* strains isolated from infant botulism, and identified as *C. butyricum* were shown to synthesize synthesize BoNT/E (10, 11). A comparison of purified BoNT/E derived from *C. botulinum* and *C. butyricum* showed that both toxins are very similar. Their molecular weights (145 000 Da) determined by polyacrylamide gel electrophoresis were in good agreement with that calculated from the deduced amino acid sequences. In addition, both toxins exhibited similar specific toxicity in mice (12).

Here, we have determined the complete nucleotide sequences of the BoNT/E genes from *C. botulinum* E strain Beluga, and from the toxigenic *C. butyricum* strains ATCC 43181 and ATCC 43755, and we present the deduced amino acid sequences.

MATERIALS AND METHODS

Bacterial DNA and Plasmids.*C. botulinum* type E strain Beluga, and toxigenic *C. butyricum* strains ATCC 43181 and ATCC 43755 received directely from ATCC, were grown in TGY broth (Trypticase, 30 g/l; yeast extract, 20 g/l; glucose, 5 g/l; HCl-cysteine, 0.5 g/l; pH7.2) in anaerobic conditions. Total DNA was extracted and purified as previously described (13).

PLasmid pUC 19 (Appligene, Strasbourg, France) was used for cloning in *Escherichia coli* strain TG1.

Probes and hybridization conditions. Oligonucleotides were synthesized by the phosphoramidite method using a Cyclone Miligen automated DNA synthesizer. Gene Screen Plus   filters (New England Nuclear Research Products, Du Pont Nemours, Boston, USA) were pretreated with 200 µg/ml heat-denaturated salmon sperm DNA in 1 M NaCl, 10% dextran sulfate, 0.5% SDS, 50 mM Tris HCl, pH 7.5 at 40°C and then with a 5'(^{32}P) labeled oligonucleotide (10^6 cpm/ml) in the same mixture overnight at 40°C. Filters were washed in 6 X SSC, 0.1% SDS at 40°C for two hours and exposed overnight to Fuji RX films.

Polymerase chain rea
amplified using the polyme
C. botulinum E strain Belt
volume of 100 µl containin
0.1% BSA, 100 µM dNTP,
and 2.5 U of Taq polymer
denaturated at 95°C for
consisting of denaturatic
temperature which was £
primers), and extension (20
Elemer Cetus, Emeryvil
Geneclean   (Bio 101 Inc.,

Other molecular biolog
plasmid DNA from *E. coli*
Bacteria were transformec
intestinal phosphatase were
from Pharmacia (Paris, F
termination procedure (15)
Corporation, Cleveland, US

RE

We used a synthet
sequence from positions 71
the BoNT/E gene to identif
fragment was then used to
with the help of the latter cl
was obtained. Together, the
BoNT/E gene from *C. botu*
bp) has been entered in the
contains a single open rea
sequence (1251 residues) is

Overlapping DNA 1
butyricum strains ATC(
oligonucleotides the seque
BoNT/E gene. The nucleot
butyricum strains were det
EMBL database under the a
A comparison of the
the *C. botulinum* and the *C*
64 of them in the coding re
acid changes were detect
terminal residues which rep
between *C. botulinum* E a

explain the development of

rnalization and intraneuronal

al. The heavy chains mediate

chains appear to enter target

known how the different heavy

r what is the molecular nature

er, these questions may be

on of conserved domains. The

sequenced previously (5, 6, 7

chromosomally located in C.

as established by cloning and

· strains quite different from C.

m strains isolated from infant

own to synthesize synthesize

ived from C. botulinum and C.

Their molecular weights (145

resis were in good agreement

iences. In addition, both toxins

tide sequences of the BoNT/E

n the toxigenic C. butyricum

sent the deduced amino acid

ODS

E strain Beluga, and toxigenic

received directly from ATCC

ract, 20 g/l; glucose, 5 g/l; HCl

Total DNA was extracted and

ance) was used for cloning in

leotides were synthesized by

n automated DNA synthesizer

· Research Products, Du Pont

g/ml heat-denatured salmon

SDS, 50 mM Tris HCl, pH 7.5 at

ide (10^6 cpm/ml) in the same

X SSC, 0.1% SDS at 40°C for

polymerase chain reaction amplification. C. butyricum DNA (100ng) was amplified using the polymerase chain reaction (PCR) and primers deduced from the C. botulinum E strain Beluga BoNT/E gene. The reactions were done in a total volume of 100 µl containing 10 mM Tris HCl, pH 8.3, 50 mM KCl, 4 mM $MgCl_2$, 0.1% EDTA, 100 µM dNTP, 10 mM beta-mercaptoethanol, 25 pmol of each primer, and 2.5 U of Taq polymerase (Beckman, Paris, France). Reaction mixtures were denaturated at 95°C for 2 min and then submitted to 30 subsequent cycles consisting of denaturation (20 s at 94°C), annealing (20s at hybridization temperature which was 5°C below the theoretical melting temperature of the primers), and extension (20 s at 72°C) in a DNA Thermal Cycler version 2.2 (Perking Elemer Cetus, Emeryville, USA). Amplification products were purified by Geneclean (Bio 101 Inc., La Jolla, USA) and sequenced.

Other molecular biological techniques. Ligation reactions and preparation of plasmid DNA from E. coli were conducted as described by Maniatis et al. (14). Bacteria were transformed by electroporation. T4 polynucleotide kinase and calf intestinal phosphatase were from Boehringer-Mannheim France, and other enzymes from Pharmacia (Paris, France). DNA was sequenced by the dideoxy-chain-termination procedure (15) using the Sequenase Kit (United States Biochemical Corporation, Cleveland, USA).

RESULTS AND DISCUSSION

We used a synthetic 42mer oligonucleotide representing the nucleotide sequence from positions 716 to 757 of the previously published EcoRI fragment of the BoNT/E gene to identify and clone an overlapping 1373 bp NsiI fragment. This fragment was then used to clone a ScaI fragment (1639 bp) to yield pMRP43, and with the help of the latter clone, pMRP46 containing an HindIII fragment of 2500 bp was obtained. Together, these four clones encompass the entire coding region of the BoNT/E gene from C. botulinum E strain Beluga. The nucleotide sequence (4017 bp) has been entered in the EMBL database under the accession number X62089. It contains a single open reading frame of 3753 nucleotides. The deduced amino acid sequence (1251 residues) is shown in Fig. 1.

Overlapping DNA fragments of 400 to 500 bp were amplified from C. butyricum strains ATCC 43181 and ATCC 43755 DNA by PCR using oligonucleotides the sequences of which were deduced from the C. botulinum E BoNT/E gene. The nucleotide sequences of the BoNT/E genes from the two C. butyricum strains were determined to be identical and have been entered in the EMBL database under the accession number X62088.

A comparison of the DNA sequences established for the BoNT/E genes from the C. botulinum and the C. butyricum strains revealed differences in 69 positions, 64 of them in the coding region and 5 in the 5' noncoding region. Thirty nine amino-acid changes were detected, 19 of them (48%) being located within the 198 N-terminal residues which represent 16% of the entire sequence. Amino-acid identities between C. botulinum E and C. butyricum BoNT/E sequences are 95% for the L

```
  1 M P K I N S F N Y N D P V N D R T I L Y I K P G G C Q E F Y
    - - T - - - - - - - - - - - N - - - - - - - - - - - - - - -

 31 K S F N I M K N I W I I P E R N V I G T T P Q D F H P P T S
    - - - - - - - - - - - - - - - - - - - - - I - - - - L - - - -

 61 L K N G D S S Y Y D P N Y L Q S D E E K D R F L K I V T K I
    - - - - - - - - - - - - - - - - - Q - - - K - - - - - - - -

 91 F N R I N N N L S G G I L L E E L S K A N P Y L G N D N T P
    - - - - - D - - - - R - - - - - - - - - - - - - - - - - - -

121 D N Q F H I G D A S A V E I K F S N G S Q D I L L P N V I I
    - G D - I - N - - - - - P - Q - - - - - S - - - - - - - -

151 M G A E P D L F E T N S S N I S L R N N Y M P S N H R F G S
    - - - - - - - - - - - - - - - - - - - - - - - - G - - -

181 I A I V T F S P E Y S F R F N D N C M N E F I Q D P A L T L
    - - - - - - - - - - - - - K - - S - - - - - - - - - - - - -

211 M H E L I H S L H G L Y G A K G I T T K Y T I T Q K Q N P L
    - - - - - - - - - - - - - - - - - - - - - - - - - - - - - -

241 I T N I R G T N I E E F L T F G G T D L N I I T S A Q S N D
    - - - - - - - - - - - - - - - - - - - - - - - - - - - - -

271 I Y T N L L A D Y K K I A S K L S K V Q V S N P L L N P Y K
    - - - - - - - - - - - - - - - - - - - - - - - - - - - - -

301 D V F E A K Y G L D K D A S G I Y S V N I N K F N D I F K K
    - - - - - - - - - - - - - - - - - - - - - - - - - - - - -

331 L Y S F T E F D L R T K F Q V K C R Q T Y I G Q Y K Y F K L
    - - - - - - - - - A - - - - - - - - - - - - - - - - - - -

361 S N L L N D S I Y N I S E G Y N I N N L K V N F R G Q N A N
    - - - - - - - - - - - - - - - - - - - - - - - - - - - - -

391 L N P R I I T P I T G R G L V K K I I R F C K N I V S V K G
    - - - - - - - - - - - - - - - - - - - - - - - - - - - - -

421 I R K S I C I E I N N G E L F F V A S E N S Y N D D N I N T
    - - - - - - - - - - - - - - - - - - - - - - - - - - - - -

451 P K E I D D T V T S N N N Y E N D L D Q V I L N F N S E S A
    - - - - - - - - - - - - - - - - - - - - - - - - - - - - -

481 P G L S D E K L N L T I Q N D A Y I P K Y D S N G T S D I E
    - - - - - - - - - - - - - - - - - - - - - - - - - - - - -

511 Q H D V N E L N V F F Y L D A Q K V P E G E N N V N L T S S
    - - - - - - - - - - - - - - - - - - - - - - - - - - - - -

541 I D T A L L E Q P K I Y T F F S S E F I N N V N K P V Q A A
    - - - - - - - - - - - - - - - - - - - - - - - - - - - - -

571 L F V S W I Q Q V L V D F T T E A N Q K S T V D K I A D I S
    - - - G - - - - - - - - - - - - - - - - - - - - - - - - -

601 I V V P Y I G L A L N I G N E A Q K G N F K D A L E L L G A
    - - - - - - - - - - - - - - - - - - - - - - - - - - - - -

631 G I L L E F E P E L L I P T I L V F T I K S F L G S S D N K
    - - - - - - - - - - - - - - - - - - - - - - - - - - - - -

661 N K V I K A I N N A L K E R D E K W K E V Y S F I V S N W M
    - - - - - - - - - - - - - - - - - - - - - - - - - - - - -

691 T K I N T Q F N K R K E Q M Y Q A L Q N Q V N A I K T I I E
    - - - - - - - - - - - - - - - - - - - - - - - L - A - - -

721 S K Y N S Y T L E E K N E L T N K Y D I K Q I E N E L N Q K
    - - - - - - - - - - - - - - - - - - - E - - - - - - - - -
```

Fig. 1 . Alignment of the deduced amino-acid sequences of BoNT/E derived from *C. botulinum* E strain Beluga (upper line), and *C. butyricum* strains ATCC 43181 c
ATCC 43755 (lower line). Dashes represent identical amino acids.

```
751 V S I A M N N .
    - - - - - - -

781 K L R E Y D E I
    - - - - - - -

811 L N S M V T D ]
    - - - - - - I - -

841 N K F F K R I I
    
871 N I N I N G D \

901 S Q N D Y I I Y

931 V N V N N E Y T
    - - - - - - -

961 W T F E D N R G
    - - L Q - - S -

991 I F V T I T N C
    - - - - - - -

1021 G N I H V S D N
     - - - - - - -

1051 D K E L D E T E
     - - - - - - -

1081 L Y D K E Y Y L
     - - - - - - -

1111 I R S T I L L A
     - - - - - - -

1141 L V R K N D Q V
     - - - - - - -

1171 K E K T I K I S
     - - - - - - -

1201 K N N N G N N I
     - - - - - - -

1231 H T N S N G C F
     N - - - - - F -
```

chain, and 98% for the H
polypeptides are 143 836.70 [
Da for BoNT/E from the *C. buty.*
    The deduced amino-a
sequences determined for the
C. botulinum and *C. butyricur.*
BoNT/E from *C. botulinum* E ai
the Cys residues located at p
disulfide bridge between the L ;
    Fujii *et al.* have previou:
fragment encoding the 5'-termir

130

781 K L R E Y D E N V K T Y L L N Y I I Q H G S I L G E S Q Q E
    - - - - - - - - - - - - - D - - - K - - - - - - - - - - -

811 L N S M V T D T L N N S I P F K L S S Y T D D K I L I S Y F
    - - - - - I -

841 N K F F K R I K S S S V L N M R Y K N D K Y V D T S G Y D S

871 N I N I N G D V Y K Y P T N K N Q F G I Y N D K L S E V N I
    -

901 S Q N D Y I I Y D N K Y K N F S I S F W V R I P N Y D N K I
    -

931 V N V N N E Y T I I N C M R D N N S G W K V S L N H N E I I
    -

961 W T F E D N R G I N Q K L A F N Y G N A N G I S D Y I N K W
    - - L Q - - S -

991 I F V T I T N D R L G D S K L Y I N G N L I D Q K S I L N L
    - K - - - - - -

1021 G N I H V S D N I L F K I V N C S Y T R Y I G I R Y F N I F
     -

1051 D K E L D E T E I Q T L Y S N E P N T N I L K D F W G N Y L
     - - - - - - - - - - - - - N - - - A - - - - - - - - - - -

1081 L Y D K E Y Y L L N V L K P N N F I D R R K D S T L S I N N
     - - - - - - - - - - - - - - - - - N - - T - - - - - - - -

1111 I R S T I L L A N R L Y S G I K V K I Q R V N N S S T N D N
     -

1141 L V R K N D Q V Y I N F V A S K T H L F P L Y A D T A T T N
     - - - - - - - - - - - - - - - - - L - - - - - - - - - - -

1171 K E K T I K I S S S G N R F N Q V V V M N S V G N C T M N F
     -

1201 K N N N G N N I G L L G F K A D T V V A S T W Y Y T H M R D
     -

1231 H T N S N G C F W N F I S E E H G W Q E K 1251
     N - - - - - F - - - - - - - - - - - - - -

Fig. 1 – continued

chain, and 98% for the H chain. The molecular weights of the predicted polypeptides are 143 836.70 Da for BoNT/E from *C. botulinum* E, and 143 389.33 Da for BoNT/E from the *C. butyricum* strains, respectively.

The deduced amino-acid sequences are in close agreement with partial sequences determined for the N-termini of the L and the H chains of BoNT/E from *C. botulinum* and *C. butyricum* (1). As shown by Das Gupta (1), the H chain of BoNT/E from *C. botulinum* E and *C. butyricum* begins at Lys in position 423. Thus, the Cys residues located at positions 412 and 426 are probably involved in the disulfide bridge between the L and H chains.

Fujii *et al.* have previously reported the nucleotide sequences of the EcoRI fragment encoding the 5'-terminus of the BoTN/E gene from *C. botulinum* E strains

Mashike, Iwanai, and Otaru, and from *C. butyricum* strain BL6340 (16, 17). Several differences were found between these nucleotide sequences and the corresponding sequences of *C. botulinum* E strain Beluga and *C. butyricum* strains ATCC 43181 and ATCC 43755. As reported in this study, the Beluga strain produces a BoNT/E that contains an Arg at position 177, Cys at 198, and Lys at 230. The Japanese BoNT/E reference strains contained a neurotoxin with Gly at position 177, Ser at 198, and Met at position 230. The presence of Met in this latter position was also detected in BoNT/E from the *C. butyricum* strain BL6340 (17). Taken together, these findings indicate that botulinal neurotoxins, although belonging to the same toxinotype, may exhibit minor differences in their amino acid sequences.

BoNT/E (1251 residues) is significantly shorter than TeTx (1315 residues), BoNT/A (1296 residues), BoNT/C1 (1291 residues) and BoNT/D (1276 residues). At the amino-acid level, BoNT/E from *C. botulinum* has an overall identity of 38.3% with TeTx, 44.0% with BoNT/A, 38.3% with BoNT/C1, and 32.0% with BoNT/D.

Functional domains which constitute the receptor binding sites on the H chains or which are involved in the as yet unspecified toxification process mediated by the L chains are expected to be conserved. An alignment of the amino acid sequences of the L chains of TeTx, BoNT/A, B, C1, D, and E revealed that indeed highly conserved domains are separated from each other by short variable regions composed of 10 to 30 unrelated residues (18). In this respect, it is noteworthy that only 1 out of the 19 amino acid exchanges found in the N-terminal portions of the two BoNT/E L chain sequences involves a residue that is conserved in the other neurotoxins, while 10 of the replacements reside in such variable regions. It is possible that these regions define the immunologic differences of the individual toxinotypes. This would explain why only five out of nine monoclonal antibodies raised against BoNT/E from *C. botulinum* E reacted with the neurotoxin from the *C. butyricum* strains (19).

The His-rich motif in the center of the BoNT/E L chain involving His at positions 212, 216, and 219, is conserved in the sequences of all clostridial neurotoxins analysed so far (7). It remains to be shown, wether this motif is directely involved in the toxification reaction or in the translocation process of the L chain into the cytosole. Alternatively, this motif could merely serve as an element that stabilizes the tertiary structures of the individual L chains (18).

The high level of identity between BoNT/E from *C. botulinum* and *C. butyricum* suggests a transfer of the BoNT/E gene from *C. botulinum* to originally nontoxigenic *C. butyricum* strains, and an independent evolution from each other by point mutations. A similar unexpected synthesis of BoNT/F in *C. baratii* has been reported previously (11, 20). These findings underscore the general observation that classification of *C. botulinum* strains on the basis of their neurotoxins bears little

significance with respect to strains.

1. DasGupta, B.R. (1989). Ir
   Ed.) pp 53-67 Academic I
2. Simpson L.L. (1989). In: E
   Ed.) pp 153-178 Academ
3. Ahnert-Hilger, G., Weller,
   (1989). FEBS Let. 242,24
4. Bittner, M.A., Habig, W.H.
5. Fairweather, N.F., Lyness.
   J. Bacteriol. 165, 21-27.
6. Eisel, U., Jarausch, W., Ge
   M., Habermann, E., and M
7. Binz, T., Kurazono, H., W
   J. Biol. Chem. 265,9153-9
8. Hauser, D., Eklund, M.W.,
   P., and Popoff, M.R. (199C
9. Binz, T., Kurazono, H., Po
   Krieglstein, K., Henschen,
   18, 5556.
10. Aureli, P., Fenici, L., Pas
    Hatheway, C.L. (1986). J
11. Suen, J.C., Hatheway, C
    Microbiol. 26,2191-2192
12. Gimenez, J.A., and Sugi\
13. Popoff, M.R., Guillou, J.F
    1703.
14. Maniatis, T., Fritsch, E.F.
    laboratory manual. Cold
    York.
15. Sanger, F., Nicklen, S., ai
    5463-5467.
16. Fujii, N., Kimura, K., Indo
    and Oguma, K. (1990). M
17. Fujii, N., Kimura, K., Yas
    N., and Oguma, K. (1991
18. Niemann, H. (1991). In: S
    Freer J., Eds). Academic
19. Kozaki, S., Onimaru, J., H
    59,457-459.
20. Hall, J.D., McCroskey, L.
    Microbiol. 21,654-655.

C. butyricum strains ATCC 4318[...]
Beluga strain produces a BoNT/E[...]
, and Lys at 230. The Japanese[...]
with Gly at position 177, Ser a[...]
let in this latter position was also[...]
L6340 (17). Taken together, these[...]
though belonging to the same[...]
nino acid sequences.

orter than TeTx (1315 residues)[...]
) and BoNT/D (1276 residues). A[...]
s an overall identity of 38.3% with[...]
id 32.0% with BoNT/D.

receptor binding sites on the H[...]
ified toxification process mediated[...]
An alignment of the amino acid[...]
C1, D, and E revealed that indeed[...]
ch other by short variable regions[...]
1 this respect, it is noteworthy that[...]
d in the N-terminal portions of the[...]
lue that is conserved in the other[...]
e in such variable regions. It is[...]
ogic differences of the individual[...]
ut of nine monoclonal antibodies[...]
d with the neurotoxin from the *C*[...]

BoNT/E L chain involving His at[...]
the sequences of all clostridial[...]
hown, wether this motif is directely[...]
ocation process of the L chain into[...]
serve as an element that stabilizes[...]
).
IT/E from *C. botulinum* and *C*[...]
ne from *C. botulinum* to originally[...]
ident evolution from each other by[...]
of BoNT/F in *C. baratii* has been[...]
rscore the general observation that[...]
is of their neurotoxins bears little[...]

REFERENCES

1. DasGupta, B.R. (1989). In: Botulinum Neurotoxin and Tetanus Toxin (Simpson L.L. Ed.) pp 53-67 Academic Press, San Diego.
2. Simpson L.L. (1989). In: Botulinum Neurotoxin and Tetanus Toxin (Simpson L.L. Ed.) pp 153-178 Academic Press, San Diego.
3. Ahnert-Hilger, G., Weller, U., Dauzenroth, M.E., Habermann, E., and Gratzl, M. (1989). FEBS Let. 242,245-248.
4. Bittner, M.A., Habig, W.H., and Holz, R.W. (1989). J. Neurochem. 53, 966-968.
5. Fairweather, N.F., Lyness, V.A., Pickard, D.J., Allen, G., and Thomson, R.O. (1986). J. Bacteriol. 165, 21-27.
6. Eisel, U., Jarausch, W., Goretzki, K., Henschen, A., Engels, J., Weller, U., Hudel, M., Habermann, E., and Niemann, H. (1986). EMBO J. 5,2495-2502.
7. Binz, T., Kurazono, H., Wille, M., Frevert, J., Wernars, U., and Niemann, H. (1990). J. Biol. Chem. 265,9153-9158.
8. Hauser, D., Eklund, M.W., Kurazono, H., Binz, T., Niemann, H., Gill, D.M., Boquet, P., and Popoff, M.R. (1990). Nucl. Acids Res. 18,4924.
9. Binz, T., Kurazono, H., Popoff, M.R., Eklund, M.W., Sakaguchi, G.S., Kozaki, S., Krieglstein, K., Henschen, A., Gill, D.M., and Niemann, H. (1990) Nucl. Acids Res. 18, 5556.
10. Aureli, P., Fenici, L., Pasolini, B., Gianfranceshi, M., McCroskey, L.M., and Hatheway, C.L. (1986). J. Infect. Dis. 154,207-211.
11. Suen, J.C., Hatheway, C.L., Steigerwalt, A.G., and Brenner, D.J. (1988). J. Clin. Microbiol. 26,2191-2192.
12. Gimenez, J.A., and Sugiyama, H. (1988). Infect. Immun. 56,926-929.
13. Popoff, M.R., Guillou, J.P., and Carlier, J.P. (1985). J. Gen. Microbiol. 131,1697-1703.
14. Maniatis, T., Fritsch, E.F., and Sambrook, J. (1982). Molecular cloning: a laboratory manual. Cold Spring Harbor Laboratory, Cold Spring Harbor, New York.
15. Sanger, F., Nicklen, S., and Coulson, A.R. (1977). Proc. Natl. Acad. Sci. USA 74, 5463-5467.
16. Fujii, N., Kimura, K., Indoh, T., Murakami, T., Yashiki, T., Tsuzuki, K., Yokosawa, N., and Oguma, K. (1990). Microbiol. Immunol. 34,1041-1047.
17. Fujii, N., Kimura, K., Yashiki, T., Indoh, T., Murakami, T., Tsuzuki, K., Yokosawa, N., and Oguma, K. (1991). J. Gen. Microbiol. 137,519-525.
18. Niemann, H. (1991). In: Sourcebook of Bacterial Protein Toxins (Alouf J., and Freer J., Eds). Academic Press. New York.
19. Kozaki, S., Onimaru, J., Kamata, Y., and Sakaguchi, G. (1991). Infect. Immun. 59,457-459.
20. Hall, J.D., McCroskey, L.M., Pincomb, B.J., and Hatheway, C.L. (1985). J. Clin. Microbiol. 21,654-655.

113

CS   Department of Microbiology, AFRC Institute of Food Research, Reading, UK.
SO   FEMS MICROBIOLOGY LETTERS, (1992 Sep 15) 75 (2-3) 225-30.
      Journal code: 7705721. ISSN: 0378-1097.
CY   Netherlands
DT   Journal; Article; (JOURNAL ARTICLE)
LA   English
FS   Priority Journals
OS   GENBANK-M92906
EM   199211
ED   Entered STN: 19930122
      Last Updated on STN: 19970203
      Entered Medline: 19921119

L1   ANSWER 16 OF 17   MEDLINE
AN   90262533   MEDLINE
DN   90262533   PubMed ID: 2188647
TI   Botulinum type F neurotoxin. Large-scale purification
      and characterization of its binding to rat cerebrocortical synaptosomes.
AU   Wadsworth J D; Desai M; Tranter H S; King H J; Hambleton P; Melling J;
      Dolly J O; Shone C C
CS   Department of Biochemistry, Imperial College of Science and Technology,
      London, U.K.
SO   BIOCHEMICAL JOURNAL, (1990 May 15) 268 (1) 123-8.
      Journal code: 2984726R. ISSN: 0264-6021.
CY   ENGLAND: United Kingdom
DT   Journal; Article; (JOURNAL ARTICLE)
LA   English
FS   Priority Journals
EM   199006
ED   Entered STN: 19900720
      Last Updated on STN: 19970203
      Entered Medline: 19900628

L1   ANSWER 17 OF 17   MEDLINE
AN   84018414   MEDLINE
DN   84018414   PubMed ID: 6353671
TI   Amino acid composition of Clostridium botulinum type F
      neurotoxin.
AU   DasGupta B E; Rasmussen S
NC   NS 17742 (NINDS)
SO   TOXICON, (1983) 21 (4) 566-9.
      Journal code: 1307333. ISSN: 0041-0101.
CY   ENGLAND: United Kingdom
DT   Journal; Article; (JOURNAL ARTICLE)
LA   English
FS   Priority Journals
EM   198311
ED   Entered STN: 19900319
      Last Updated on STN: 19970203
      Entered Medline: 19831123

Thanks,
Sharon L. Turner, Ph.D.
USPTO
CM1-10B09
Mailroom 10C01
Biotechnology GAU 1647
(703) 308-0056

., (1982) Nu-

old, B.S. and

Annu. Rev.

nan-Hansen,

and Collins,
132.
E. (1988) J.

, T. (1989)
d edn. Cold
ing Harbor,

Appl. Bacte-

Natl. Acad.

(1984) Nu-

. and Noller,

1982) Gene

Sequence of the gene encoding type F neurotoxin of *Clostridium botulinum*

Alison K. East, P.T. Richardson, D. Allaway, M.D. Collins, T.A. Roberts
and Daphne E. Thompson

Department of Microbiology, AFRC Institute of Food Research, Reading Laboratory, Reading, UK

1. SUMMARY

Primers designed to conserved regions of botulinum and tetanus clostridial toxins were used to amplify DNA fragments from non-proteolytic *Clostridium botulinum* type F (202F) DNA using polymerase chain reaction technology. The fragments were cloned and the complete nucleotide sequence of the gene encoding type F toxin determined. Analysis of the nucleotide sequence demonstrated the presence of an open frame encoding a protein of 1274 amino acids, similar to other botulinum neurotoxins. Upstream of the toxin gene is the end of an open reading frame which encodes the C-terminus of a protein with homology to non-toxic-non-hemagglutinin component of type C progenitor toxin.

Correspondence to: D.E. Thompson, Department of Health, Eileen House, 80-94 Newington Causeway, London, SE1 6EF, UK.

2. INTRODUCTION

Within the genus *Clostridium*, organisms producing protein toxins able to cause paralysis symptomatic of botulism are deemed to be *Clostridium botulinum*. Outbreaks of botulism in humans and animals have resulted in the classification of the neurotoxin into seven serologically different forms identified as type A (BoNT/A) to type G (BoNT/G). BoNT/A, BoNT/B and BoNT/E are the main causative agents of human botulism, with rare incidents attributed to BoNT/F and BoNT/G [1]. The disease is most often the result of ingestion of toxin formed in foods that have been inadequately processed and/or stored at inappropriate temperatures. Another source of the disease is after colonization of the gut by clostridia in infants of less than approx. 6 months, consequently known as infant botulism [1,2]. Occasionally botulism is the result of infection of a wound. There have been reports of BoNT being produced by clostridia that

226

phenotypically are not *C. botulinum* viz. *C. butyricum* and *C. barati* [1,2]. Type F neurotoxin is produced by *C. botulinum* Group I (proteolytic strains), Group II (non-proteolytic strains) and by *C. barati* [3].

Comparison of the complete amino acid sequences from BoNT genes (BoNT/A [4,5], BoNT/C [6], BoNT/D [7], BoNT/E [8] and BoNT/E of *C. butyricum* [9]), and partial sequence (BoNT/B [10]), with tetanus toxin (TeTx) [11] has revealed conserved regions thought to be implicated in toxin function [10]. Primers based on data from these conserved regions were used to clone and sequence the gene encoding BoNT/F from non-proteolytic *C. botulinum* using polymerase chain reaction (PCR). We report here the nucleotide sequence of a gene, which when translated, gives an amino acid sequence similar to those of BoNTs already reported [5–10].

3. MATERIALS AND METHODS

3.1. Cultures and preparation of DNA

C. botulinum strain 202F (ATCC 23387) was grown anaerobically in TPYCG (trypticase 50 g/l, bactopeptone 5 g/l, yeast extract 3 g/l, cysteine HCl 0.5 g/l and glucose 5 g/l) broth and DNA prepared as described by Farrow et al. [12]. All manipulations of *C. botulinum* and cloned fragments were carried out under GMP II conditions. *E. coli* SURE (Stratagene) was used as the recipient for transformation. Plasmid DNA was prepared from *E. coli* by the alkaline lysis method of Birnboim and Doly (see ref. 13) and for sequencing, by the method of Treisman as described by Sambrook et al. [13].

3.2. PCR amplification

PCR was performed with a Biometra thermocycler using the following conditions: 95°C for 5 min, followed by 25 cycles of: 94°C for 45 s, between 37 and 45°C for 45 s and 58°C for 2 min, followed by 62°C for 10 min and then 4°C. Primers were made on an Applied Biosystems model 391 DNA synthesiser and used at a final concentration of 4 ng/μl. Deoxynucleotides (BCL) were present at a final concentration of 200 mM and 1

unit Taq polymerase (Amersham International) was used according to the manufacturers instructions. Template DNA was at a final concentration of 20 ng/μl except for 'inverse PCR' when the concentration was 50 ng/μl of restriction enzyme-digested and re-ligated DNA. For 'inverse PCR' [14], DNA was prepared using restriction enzymes (BCL) according to the manufacturer's instructions and ligations performed using a 'Ligation System' kit (Amersham International).

3.3. Cloning and transformation

Products of PCR were cloned into the *Eco*RV site of pBluescript KS + (Stratagene) following modification and purification of the fragment by Klenow (BCL), 'Geneclean' (Stratatech), kinase (Gibco-BRL), phenol extraction and ethanol precipitation [13]. An alternative vector system was used to construct pCBF2 and pCBF8; DNA fragments generated by PCR were cloned directly, without enzymic modification, into 'T-vector' (D. Evans, University of Reading, personal communication) digested with *Xcm*I (New England Biolabs). This digestion generates a single 3' overhanging T on the vector which base pairs with the A, often added to the 3' end of PCR fragments during the reaction [15]. Transformation into *E. coli* competent cells, prepared as described by Inoue et al. [16], was carried out as described by Sambrook et al. [13]. Plasmids containing inserts of the expected size were sequenced following testing of the recombinant-cell lysates by the mouse bioassay [17], to ensure that no toxic polypeptides were produced.

3.4. Sequence determination

Sequencing of plasmid DNA was carried out using dideoxy-nucleotide chain termination. Primers complementary to both the plasmid and toxin were used to sequence both strands of each insert completely. As the inserts were generated using PCR, the sequence of two fragments synthesised in different PCR reactions was determined. Where two sequences differed, as they did in three positions, a third clone was sequenced and in each case confirmed one of the sequences.

4. RESU

To e
produce
BoNT/F
the posi
gions of
No reco
bioassay
The c
1274 an
146 708.
the mea
fied fro
tein enc
with Te
are usu
have dif
proteins
for con
with lig
sequenc
chain of
showing
of C. b
than Tc
identity
homolo
(64.9%
tity), w
44.6% a
identity

Sau3Al

Fig. 1. Sc
fragment
digested

national)
s instruc-
oncentra-
:R' when
estriction
For 'in-
g restric-
manufac-
ied using
iational).

: EcoRV
following
ment by
), kinase
inol pre-
tem was
NA frag-
directly,
ctor' (D.
ommuni-
ind Bio-
3' over-
with the
agments
into E.
ibed by
ribed by
; inserts
ollowing
by the
to toxic

ied out
ination.
nid and
of each
nerated
nts syn-
; deter-
hey did
luenced
uences.

4. RESULTS AND DISCUSSION

To ensure that no toxic polypeptides were produced only fragments encoding part of BoNT/F were cloned in *E. coli*. Figure 1 shows the position of the primers used to amplify regions of DNA which were subsequently cloned. No recombinant plasmid was toxic in the mouse bioassay [17].

The open reading frame codes for a protein of 1274 amino acids with a molecular mass of 146 708. This calculated value compares well with the measured value of 150 000 for BoNT/F purified from proteolytic *C. botulinum* [18]. The protein encoded by BoNT/F shows high homology with TeTx and other BoNTs. Since these toxins are usually cleaved *in vivo* and as the two chains have different functions in toxic activity [10], the proteins were divided into light and heavy chains for comparative purposes. Comparing BoNT/F with light and heavy chains for toxins where the sequence is known (data not shown), the light chain of BoNT/F is more like that of BoNT/E, showing 58.3% and 57.6% identity with the toxins of *C. botulinum* and *C. butyricum* respectively, than TeTx (45.1% identity) and BoNT/A (35.1% identity). The heavy chain also showed greater homology with the BoNT/E from *C. botulinum* (64.9% identity) and *C. butyricum* (64.4% identity), with BoNT/A and TeTx having identities 44.6% and 35.0% respectively. The high sequence identity of BoNT/F and BoNT/E may explain

the cross-reactivity of antisera raised against either type toxin with both BoNT/F and BoNT/E.

Within the toxin sequence there are regions of highly conserved amino acids thought to be important for structure and function (see ref. 10). In the light chain, the region noted by Niemann, DPhhnLhHELnHnnHxLYG, where h is a hydrophobic residue, n is uncharged and x is any amino acid [10], is also present in BoNT/F (Fig. 2, residues 220–239). The cysteine residues involved in disulphide linkage of the heavy and light chains are found in all the toxins (Fig. 2, residues 429 and 445). The number of residues between the cysteines varies between eight (BoNT/B) and 27 (TeTx); BoNT/F has 15 amino acids. As with BoNT/E, this BoNT/F is encoded by a non-proteolytic *C. botulinum*, it is therefore unlikely to undergo proteolytic cleavage by an endogenous protease, but a possible cleavage site for exogenous proteases is between residues 439 and 440 (Fig. 2). The C-terminus of the heavy chain, a region noted for its lack of homology in BoNTs [10], is very well conserved between BoNT/F and BoNT/E. It has been suggested that this region is involved in the binding of BoNT to a ganglioside receptor [10]. The high homology between BoNT/E and BoNT/F may indicate that they share a common receptor. Analysis of sequences upstream of BoNT/F revealed a possible Shine-Dalgarno sequence 5'-AGGGGG-3' (Fig. 2, nucleotides 132–137), eight nucleotides upstream of the proposed start of
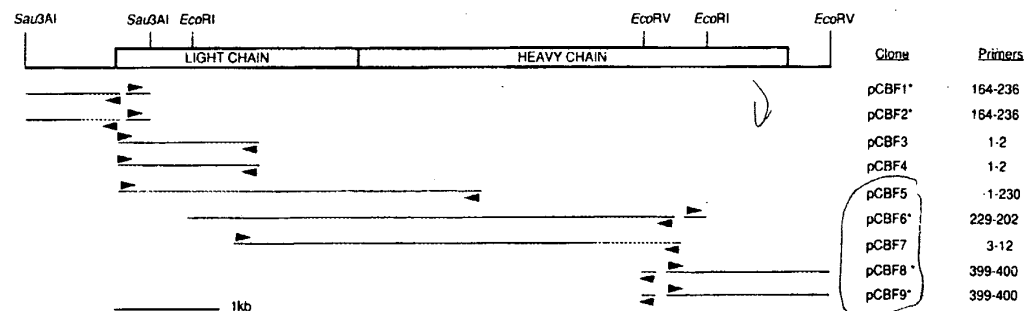
Fig. 1. Strategy for cloning fragments of the BoNT/F gene. Arrows indicate the position of the primers used to generate the fragments. Clones generated by 'inverse PCR' are indicated by *; in the case of pCBF1*/pCBF2* the chromosomal DNA was digested with *Sau*3AI, for pCBF6* with *Eco*RI, and pCBF8*/pCBF9* with *Eco*RV. Only the sites of restriction enzymes involved in the construction of clones are shown.

translation. Further upstream, nucleotides 33–38 and 47 (Fig. 2) are identical to those mapped as the −10 and transcriptional start by Binz et al. for BoNT/A using primer extension [5]. The −35 region suggested by the same authors for BoNT/A of 5'-TTAACCC-3' is present virtually unchanged (5'-TTAACGC-3') in the same position in BoNT/F (Fig. 2, nucleotides 5–11). However, further analysis of this upstream region of the BoNT/F gene shows that it contains the 3'-end of an open reading frame (Fig. 2, nu-

cleotides 1–130 and data not shown), translation of which into protein sequence gives the C-terminal 57 amino acids (Fig. 3). The translated sequence shows homology with the non-toxic-non-haemagglutinin component of C. botulinum type C [19], the gene for which is found in the same position, upstream of the toxin gene. Published sequences reveal the presence of such open reading frames in type A [5], type E [8] and C. butyricum [9], although not commented upon by the authors. As these translated open reading

Fig. 2. Complete nucleotide sequence of the BoNT/F gene. The translated amino acid sequence of BoNT/F is given under the second nucleotide of each codon. The position of primers used in PCR to generate the clones is indicated by underlining; * above a nucleotide indicates a mismatch between the primer and the DNA sequence. The possible position of cleavage between the light and heavy chains is indicated. The cysteine residues are underlined at amino acids 364 and 445. Nucleotides 1 to 130 form part of the open reading frame translated to give the sequence shown in Fig. 3. The sequence has been given the Genbank database accession number M92906.

translation
e C-termi-
slated se-
toxic-non-
inum type
the same
Published
pen read-
] and *C.*
l upon by
n reading

ATAAAAACAGCAATAGAATATAAATATAATAATAATTATACTTCAGATGAGAAAAATAGACTT 2400
I K T A I E Y K Y N K Y T S D E K N R L (752)

GAATCTGAATATAATATCAATAATATAGAAGAAGAATTGAATAAAAAAGTTTCTTTAGCA 2460
E S E Y N I N N I E E E L N K K V S L A (772)

ATGAAAAATATAGAAAGATTTATGACAGAAAGTTCTATATCTTATTTAATGAAATTAATA 2520
M K N I E R F M T E S S I S Y L M K L I (792)

AATGAAGCCAAAGTTGGTAAATTAAAAAAAATATGATAACCATGTTAAGAGCGATTTATTA 2580
N E A K V G K L K K Y C N H V K S D L I (812)

AACTATATTCTCGACCATAGATCAATCTTAGGAGAGCAGACAAATGAATTAAGTGATTTG 2640
N Y I L D H R S I L G E Q T N E L S C L (832)

GTGACTAGTACTTTGAATAGTAGTATTCCATTTGAACTTTCTTCATATACTAATGATAAA 2700
V T S T L N S S I F F E L S S Y T N D K (852)

ATTCTAATTATATACTTTAATAGATTTATATAAAAAAATTAAAGATAGTTCTATTTTAGAT 2760
I L I I Y F N R L Y K S I K D S S I L D (872)

ATGCGATATGAAAATAATAAATTTATAGATATCTCTGGATATGGTTCAAATATAAGCATT 2820
M R Y E N N K F I C I S G Y G S N I S I (892)

AATGGAAACGTATATATTTATTCAACAAATACAAATCAATTTGGAATATATAATAGTAGG 2850
N G N V Y I Y S T N R N Q F G I Y N S R (912)

CTTAGTGAAGTTAATATAGCTCAAAATAATGATATTATATACAATAGTAGATATCAAAAT 2940
L S E V N I A Q N N C I I Y N S R Y Q N (932)

TTTAGTATTAGTTTCTGGGTAAGGATTCCTAAACACTACAAACCTATGAATCATAATCGG 3000
F S I S F W V R I P K H Y K P M N H N R (952)

GAATACACTATAATAAATTGTATCGGGAATAATAATTCGGGATGGAAAATATCACTTAGA 3060
E Y T I I N C M G N N N S G W K I S L R (972)

ACTGTTAGAGATTGTGAAATAATTTGGACTTTACAAGATACTTCTGGAAATAAGGAAAAT 3120
T V R D C E I I W T L Q D T S G N K E N (992)

TTAATTTTTAGGTATGAAGAACTTAATAGGATATCTAATTATAATAAATAAATGGATTTTT 3180
L I F R Y E E L N R I S N Y I N K W I F (1012)

GTAACTATTACTAATAATAATAGATTAGGCAATTCTAGAATTTACATCAATGGAAATTTAATA 3240
V T I T N N R L G N S R I Y I N G N L I (1032)

GTTGAAAAATCAATTTCGAATTTAGGTGATATTCATGTTAGTGATAATAATAGTTATTTAAA 3300
V E K S I S N L G D I H V S D N I L F K (1052)

ATTGTTGGTTGTGATGATGAAACGTATGTTGGTATAAGATATTTTAAAGTTTTTAATACG 3360
I V G C D D E T Y V G I R Y F K V F N I (1072)

GAATTAGATAAAACAGAAATTGAGACTTTATATAGTAATGAGCCAGATCCAAGTATCTTA 3420
E L D K T E I E T L Y S N E F D P S I L (1092)

AAAAACTATTGGGGAAATTATTTGCTATATAATAAAAAAATATTATTTATTCAACTTACTA 3480
K N Y W G N Y L L Y N K K Y Y L F N L L (1112)

Fig. 2. (continued).

EcoRI
AGAAAAGATAAGTATATTACTCTGAATTCAGGCATTTTAAATATTAATCAACAAAGAGGT 3540
R K D K Y I T L N S G I L N I V Q Q R G (1132)

GTTACTGAAGGCTCTGTTTTTTTGAACTATAAATTATATGAAGGAGTAGAAGTCATTATA 3600
V T E G S V F L N Y K L Y E G V E V I I (1152)

AGAAAAAATGGTCCTATAGATATATCTAATACAGATAATTTTGTTAGAAAAAACGATCTA 3660
R K N G P I D I S N T D N F V R K N D L (1172)

GCATACATTAATGTAGTAGATCGTGGTGTAGAATATCGGTTATATGCTGATACAAAATCA 3720
A Y I N V V D R G V E Y R L Y A D T K S (1192)

GAGAAAGAGAAAATAATAAGAACATCTAATCTAAACGATAGCTTAGGTCAAATTATAGTT 3780
E K E K I I R T S N L N D S L G Q I I V (1212)

ATGGATTCAATAGGAAATAATTGCACAATGAATTTTCAAAACAATAACGGGAGCAATATA 3840
M D S I G N N C T M N F Q N N N G S N I (1232)

GGATTACTAGGTTTTCATCCAAATAATTTGGTTGCTAGTAGTTGGTATTATAACAATATA 3900
G L L G F H N N L V A S S W Y Y N N I (1252)

CGAAGAAATACTAGCAGTAATGGATGCTTTTGGAGTTCTATTTCTAAAGAGAATGGATGG 3960
R R N T S S N G C F W S S I S K E N G W (1272)

AAAGAATGATAGGCCACAGTAATTAATCTCAAACTACATGAGTCTGTCAAGAATTTTTGT 4020
K E (1274)

ATAAACATTCATAAAAAAATTTTAAAAATTAAIATATTTAACAATAGCTAGATATGAGTATT 4080

GCTATGTTAATATCTAGCTATTTTAATTTATCCCATATTATTATAATAAGAAAAAATACT 4140
EcoRV
ATTTCTACTGTAAGTACAAGTTTATTGGTATCTCATAAATGATACAAGATATC 4192

---

frames are homologous with the non-toxic-non-haemagglutinin component of BoNT/C (Fig. 3) it is possible that there is a toxin operon encoding the components of the progenitor botulinum toxin complex in types A and E and *C. butyricum*, as in type C [19].

Fig. 3. Comparison of the amino acid sequence of the C-terminus of open reading frames upstream of BoNT. The protein encoded by type C has been described as non-toxic-non-haemagglutinin component of the progenitor toxin [19]. For type A, type E and type E (but) the amino acid sequence was derived from published nucleotide sequences [5,8,9]. Type E (but) refers to type E of *C. butyricum*.

REFERENCES

[1] McCroskey, L.M., Hatheway, C.L. Woodruff, B.A., Greenberg, J.A. and Jurgenson, P. (1991) J. Clin. Microbiol. 29, 2618–2620.

[2] McCroskey, L.M., Hatheway, C.L., Fenicia, L., Pasolini, B. and Aureli, P. (1986) J. Clin. Microbiol. 23, 201–202.

[3] Hatheway, C.L. (1989) In: Botulinum neurotoxin and tetanus toxin, (Simpson, L.L., Ed.), pp. 3–24. Academic Press, New York.

[4] Thompson, D.E., Brehm, J.K., Oultram, J.D., Swinfield, T.-J., Shone, C.C., Atkinson, T., Melling, J. and Minton, P. (1990) Eur. J. Biochem. 189, 3–81.

[5] Binz, T., Kurazono, H., Wille, M., Frevert, J., Wernars, K. and Niemann, H. (1990) J. Biol. Chem. 265, 9153–9158.

[6] Hauser, D., Eklund, M.W., Kurazono, H., Binz, T., Niemann, N., Gill, D.M., Boquet, P. and Popoff, M.R. (1990) Nucleic Acids Res. 18, 4924.

[7] Binz, T., Kurazono, H., Popoff, M.R., Eklund, M.W., Sakaguchi, G., Kozaki, S., Krieglstein, K., Henschen, A., Gill, D.M. and Niemann, H. (1990) Nucleic Acids Res. 18, 5556.

[8] Whelan, S.M., Elmore, M.J., Bodsworth, N.J., Atkinson, T. and Minton, N.P. (1992) Eur. J. Biochem. 204, 657–667.

[9] Poulet, S., Hauser, D., Quanz, M., Niemann, H. and Popoff, M.R. (1992) Biochem. Biophys. Res. Commun. 183, 107–113.

[10] Niemann, H. (1991) In: Sourcebook of Bacterial Protein Toxins (Alouf, J.E. and Freer, J.H., Eds.), pp. 303–348. Academic Press, London.

[11] Fairweather, N.F. and Lyness, V.A. (1986) Nucleic Acids Res. 14, 7809–7812.

[12] Farrow, J.A.E., Jones, D., Phillips, B.A. and Collins, M.D. (1988) J. Gen. Microbiol. 129, 1423–1432.

[13] Sambrook, J., Fritsch, E.F. and Maniatis, T. (1989) In: Molecular cloning: A laboratory manual, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY.

[14] Ochman, H., Gerber, A.S. and Hartl, D.L. (1988) Genetics, 120, 621–623.

[15] Clark, J.M. (1988) Nucleic Acids Res. 16, 9677–9686.

[16] Inoue, H., Nojima, H. and Okayama, H. (1990) Gene 96, 23–28.

[17] Hauschild, A.H.W. (1989) In: Foodborne bacterial pathogens, (Doyle, M.P., Ed.), pp. 111–189. Marcel Dekker, New York.

[18] Yang, K.H. and Sugiyama, H. (1975) Appl. Microbiol. 29, 598–603.

[19] Tsuzuki, K., Kimura, K., Fujii, N., Yokosawa, N. and Oguma, K. (1992) Biochem. Biophys. Res. Commun. 183, 1273–1279.